DOCUMENT RESUME

ED 118 097 IR 002 998

TITLE	The Library of Congress as the National Bibliographic Center.
INSTITUTION	Association of Research Libraries, Washington, D.C.
PUB DATE	Feb 76
NOTE	63p.
AVAILABLE FROM	Association of Research Libraries, 1527 New Hampshire Avenue, N.W., Washington, D.C. 20036 ($4.00)
EDRS PRICE	MF-$0.83 HC-$3.50 Plus Postage
DESCRIPTORS	*Cataloging; Conference Reports; Data Bases; Information Processing; Information Systems; International Programs; *Library Automation; Library Planning; Library Technical Processes; *National Libraries; *National Programs; Serials; Standards; Union Catalogs
IDENTIFIERS	*Library of Congress; Machine Readable Cataloging; MARC

ABSTRACT

In a program sponsored by the Association of Research Libraries (ARL), the Library of Congress (LC) Processing Department reports on progress in development of the National Bibliographic Service. The first paper deals with past, present, and future cooperation between LC and the ARL. The second paper provides an overview of the automation activities of LC's processing department and serves as an introduction to the three subsequent papers. In the first of these the LC core bibliographic system is described, with heavy emphasis on MARC systems operations. The next paper covers the national bibliographic system including: CONSER, a serials data base; and automated register of additional locations; international cooperation; and participation in the projects of various library organizations. Finally, various problems encountered in transition to an automated system are described. The informal discussion held at the end of the program is also transcribed. (LS)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED118097

THE LIBRARY OF CONGRESS AS THE NATIONAL BIBLIOGRAPHIC CENTER

Report of a program sponsored by the

Association of Research Libraries

October 16, 1975

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ASSOCIATION OF RESEARCH LIBRARIES

FEBRUARY 1976

IR 002 998

2

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY

Association of
Research Libraries

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

Copyright © 1976 Association of Research Libraries

FOREWORD

Almost exactly seven years ago the Association of Research Libraries held its midwinter meeting in Washington, D. C. A special feature of that meeting was a day-long session at the Library of Congress designed to bring ARL directors up to date on LC activities and plans, in particular those of the Reference and Processing Departments. As a member of the group on that Saturday in January of 1969, I remember the enthusiasm with which I anticipated the day's events. I was not disappointed.

There were presentations of the work of the Reference Department by the department head and other key officials. There were tours of the Processing Department with on-site explanations of the activities of the various sections. And there was throughout an eagerness on the part of the LC staff to share its plans for the future and to secure the reactions of an important part of the Library's national clientele. The occasion was one of value to me both as the director of a rapidly-developing university library and as a newly-elected member of the ARL Board of Directors. In the first instance the presentation acquainted me with the full range of LC services available to other libraries, and in the second instance it helped me to understand the inseparable relationship between the Library of Congress and the other research libraries of the continent--especially those libraries constituting the ARL.

That day at LC left a strong impression of time well spent and useful information exchanged. It is not surprising therefore, that when planning began for the October 1975 meeting of the ARL in Washington, I remembered the success of the earlier occasion and suggested to outgoing President Richard De Gennaro and incoming President Virginia Whitney that we explore LC's interest in again sponsoring an ARL visit. Indeed, such a visit was arranged, including tours of the Processing Department and other parts of the Library, but much more was to develop from that initial suggestion. An entire program took shape, a program setting out in substantial detail important plans for the future role of the Library of Congress as the tional bibliographic center.

The substance of that program is contained in the pages that follow. Although the same material will appear later in the ARL Minutes we have undertaken to bring out this separate publication in order to share with a wider audience plans of great potential benefit to libraries everywhere.

John P. McDonald
Executive Director
Association of Research Libraries

January 1976

TABLE OF CONTENTS

	Page
FOREWORD John P. McDonald	i
LIST OF PARTICIPANTS	1
THE LIBRARY OF CONGRESS AND THE ARL Warren J. Haas	2
AUTOMATION ACTIVITIES OF THE PROCESSING DEPARTMENT William J. Welsh Henriette Avram	7
THE CORE BIBLIOGRAPHIC SYSTEM Lucia Rather	14
THE NATIONAL BIBLIOGRAPHIC SYSTEM Henriette Avram	35
TRANSITION TO THE AUTOMATED SYSTEM John C. Rather	49
DISCUSSION	53

LIST OF PARTICIPANTS

Henriette D. Avram
Chief, MARC Development Office, The Library of Congress

Richard De Gennaro
Director, University of Pennsylvania Libraries

Warren J. Haas
Vice President & Librarian, Columbia University Libraries

John J. Rather
Chief, Technical Processes Research Office, The Library of Congress

Lucia J. Rather
Assistant Chief, MARC Development Office, The Library of Congress

William J. Welsh
Director, LC Processing Department

THE LIBRARY OF CONGRESS AND THE ARL

Warren J. Haas

MR. DE GENNARO: The evening before our last membership meeting, which was held in Houston in May, we had an informal program which was called "Getting To Know The ARL." The idea behind that program was that ARL was growing and developing at the same time that its membership was increasing and many new directors were coming in as a result of a surge of retirements and new appointments. We felt it was time to reintroduce the members to ARL and ARL to the members. The success of that program inspired us to plan another introductory type program for this meeting.

For quite some time we have felt the need for an ARL program focusing on the Library of Congress, and particularly on its role as the national bibliographic center. Like ARL, LC has also been developing new roles and changing. Many of our new members need to be introduced to the Library of Congress and its work, and some of our old members need to become reacquainted with it. By a stroke of good fortune, the LC Processing Department staff under William Welsh, was primed and ready to put on a major program for us at the very time that this meeting was scheduled to be held here in Washington. As it turns out, we also have the added and unexpected pleasure of being able to meet and welcome Daniel Boorstin, the new Librarian of Congress, at this meeting featuring the national bibliographic functions of the Library of Congress.

The program is off to an excellent start. Many of you have already toured the Library of Congress this afternoon and we are all looking forward to the LC program tomorrow morning. To help us get into the right frame of mind for that program and to help us to get a better perspective on the relationship of ARL to the Library of Congress-- past, present, and future--the Board prevailed upon Warren Haas, Vice President for Information Services and University Librarian at Columbia and a tireless worker for ARL, to speak to us briefly this evening on the subject of ARL and LC. It is my pleasure to introduce Warren Haas.

* * * *

MR. HAAS: I think I will begin these remarks with what might well be my conclusion. ARL as an organization, and in this I think I can speak for its component parts as well, owes a substantial debt of gratitude to its most prominent member, the Library of Congress. Taken together, the many distinctive and important programs undertaken by LC, independently and in concert with ARL, have helped individual research libraries enormously as they have sought to fulfill their obligations to the country's students and scholars. The results of LC's efforts and expenditures have been amplified hundreds of times over in academic and research

libraries across the country.

Tomorrow morning's program, during which William Welsh and his colleagues will describe their plans for the ever-expanding set of programs that have more and more become an integral part of our own operations, promises to be a kind of high-water mark in the LC/ARL relationship that began, somewhat haltingly, about 40 years ago.

Frank McGowan, in a chapter of his 1972 thesis, The Association of Research Libraries, 1932-1962, notes that despite an early ARL/LC collaborative effort, that of creating an interlibrary loan clearinghouse through the LC Union Catalog, participation by LC in ARL was at best low-keyed during the first four or five years of ARL's history, a situation that prompted a note from Louis Hanke (then the LC-designated representative to ARL) to Archibald MacLeish suggesting that the time had arrived for the Library of Congress to play its "proper part" in the Association. The spirit of the preceding years prevailed, however, and there was no response. But a dramatic success was not long in coming.

In 1940, Paul North Rice, following up on discussions within ARL and acting in his capacity as an ARL Committee Chairman, raised with MacLeish a proposal to print the Library of Congress Catalog in book form, a suggestion that was promptly acted upon. The results were and still are clearly remarkable, since that publication, with its supplements and then its successor, the National Union Catalog, constitute a bibliographic endeavor without equal in size and complexity. Today there are 17 titles included in the most recent listing of LC catalogs in book form.

Succeeding years saw less attention given bibliographic measures, and more to resource development. During World War II, ARL and LC worked together to secure State Department approval for the pioneering Cooperative Acquisitions Project for wartime publications, an effort that brought to American research libraries large quantities of materials published abroad during the war. In 1946 ARL discussions prompted LC to establish the Documents Expediting Project, and in 1947, following several years of planning with LC and ARL, the Farmington Plan was launched. When compared with this flurry of postwar activity, the 1950's saw little in the way of additional programs, although there were several ARL/LC efforts to produce planning documents, some of great importance, on the full range of typical topics.

So far as LC/ARL relationships are concerned, it has been the last decade or so that has seen the beginning of several of the most significant projects stemming from our long association: the PL 480 Program got underway in 1962; the NPAC program in 1965; the National Serials Data Program in 1966; and the Foreign Newspaper Microfilm Project in 1970.

Since we are moving forward in time to the point where "junior citizens" like myself are fully aware of the importance and impact of

these recent projects, it is not necessary to spend time on descriptions. It is sufficient to say that the programs of the last ten years or so have established an irreversible dependency of ARL libraries on LC. The beginnings of what is essentially a bibliographic bond linking research libraries to each other in fundamental ways has now been forged, and as I see it, the work we must share during the years immediately ahead is to extend that new bond and put it to full use.

At this point, I want specifically to thank William Welsh for his commitment to the principles of a cohesive national bibliographic system. More than anyone, he has by his energy and perception breathed life into the aspirations that initially prompted NPAC, and he and his colleagues who have shared the work should know of our gratitude.

But back to our newly forged bond. Somehow its very existence carries with it substantial obligations for all parties to make it stronger and more purposeful. Perhaps with this goal in mind, it is not inappropriate to spend just a few more minutes tonight considering the future.

While the history of the ARL/LC relationship records substantial accomplishments, it also suggests that at least some of this success was perhaps too long in coming, that progress has at times been sporadic, and that the ad hoc approach to problem solving that has, until now, characterized our joint efforts has also produced less of a sense of purposeful direction than seems desirable.

The issues we face today, individually and collectively, are too important and certainly too complex for us to leave their resolution to chance. Even the most cryptic identification of these topics underscores their difficulty. By way of example, there seems to be general agreement that at least the following capacities are required:

1. There should be a national, comprehensive bibliographic data system, the components of which must assume responsibility for (a) the full range of administrative activities, (b) the on-going creation of bibliographic records, and (c) the design and operation of the systems required to distribute and use the records. The need to resolve many secondary but critical issues (e.g. standardization and international coordination) is implicit.

2. Because of the magnitude of the task, the subject of conversion of existing records to machine-readable form needs to be separately addressed, but there seems little doubt that a strategy for conversion and a program of action to accomplish the work is required.

3. Linked to improved bibliographic control is a requirement for improved access to resources. One without the other is meaningless. Further, the prospect of easily available and far more sophisticated approaches to identifying and locating information will in all probability have a dramatic effect on the level of demand for information.

4. Finally, the future nature of our collections - their content, their format, their preservation, and their distribution - raises many issues that must be resolved. The nature of that resolution will have substantial impact on the character of every library represented here.

Coupled with the magnitude of these specific items are three overriding factors that are most powerful in both pressing for and opening up the way to change. These are: (a) the potential of technology, (b) the limits of financial resources, and (c) the expanding expectations of both those who use and those who operate research libraries.

There has been a great deal of imaginative effort expended in recent years on both the basic topics and on the forces that have created a climate for change. But despite all of this accumulated wisdom, we still seem to lack a capacity for cohesive action of the kind that is required if major transformation of the research library structure of the country is to be accomplished.

Perhaps the time has now come to focus our attention on our commonly held responsibilities as much as on our specific local problems. If we are to meet in a fiscally responsible way our expanding obligations, fundamental change in the structure of research libraries and in the entire system of scholarly communication is required. This country needs a comprehensive system for the bibliographic control of all recorded information; it needs assured access to required information; it needs imaginative use of the technology that now seems capable of making a dramatic transformation possible and, finally, it needs assurance that collections of true distinction, along with related specialist staffs, are being maintained and developed, because without this assurance the importance of all of our other efforts is degraded.

We cannot assume each others service responsibilities, but by our very nature we share certain obligations that can be met only through a true partnership. The magnitude of the job to be done is such that it will require all of our best efforts. We are not part of an information industry. We are the institutions, public and private, to which society has assigned responsibility for an important set of objectives relating to assembling, preserving, and promoting the use of recorded knowledge.

No one has a precise blueprint describing where to go from here, but it does seem certain that the Library of Congress, as the country's leading research library, and the Association of Research Libraries, as the principal organization concerned specifically with the library role in support of scholarship, have essential roles to play. Perhaps the Library of Congress should take the lead by formally establishing a durable, formal and comprehensive planning effort targeted on the basic issues and the reasonably obvious objectives central to research library service and operating with the full participation of the research library community. Without specific attention to planning, especially in the area of bibliographic control, we run a substantial risk of perpetually

refining past practices that might prove unacceptable for the future. We need to establish once and for all the position of the Library of Congress as the focal point for the relatively small number of national programs that can serve as the base for a transformation in the character and capacities of research libraries.

For the ARL's part, each member has something to offer as a contributing partner to the development and implementation of national programs, and ways must be found to exploit that capacity. As an association, there is room for improvement in the ways we support the Library of Congress in its search for guidance, for funding, and for public comprehension. We would hope that the Library would turn more often to ARL for this support.

In many ways, this country is looking at 1976 not only as a historical milestone, but also as a checkpoint along a continuing path. Perhaps we can begin something now that will enable us, in the year 1982 and on the 50th anniversary of ARL, to look about and see a sophisticated, reliable, effective and financially viable system for the identification and distribution of recorded information -- a system in which technology is effectively employed, social obligations are met, and public comprehension of our goals and efforts clearly established. I think we know where we want to go -- our problem is to discover how.

* * * *

AUTOMATION ACTIVITIES AT THE LIBRARY OF CONGRESS

William J. Welsh
Director, LC Processing Department

Henriette D. Avram
Chief, MARC Development Office

MR. WELSH: I am going to give a brief introduction which will consist of an update of a statement I made before the Council for Computerized Library Networks. It has been revised to reflect many of the comments that have been made about the statement.

During the past 75 years through its service to the general library community, the Library of Congress has become, de facto, the national library in every sense of the term. Thus, in view of its demonstrated performance, its unparalleled resources, and unmatched expertise, the Library should continue to serve as the national center as we move into a new era of bibliographic control made possible by computerized library networks. The Library's role will be to develop and maintain standard bibliographic devices that will promote consistency in decentralized input to a comprehensive national data base. Decentralized input is a requirement for a national system because the Library of Congress recognizes that it cannot supply 100 percent of the cataloging information that is required nationally. Inevitably, the Library will fall short of total coverage because it will never acquire some bibliographic items; for example, many state and local documents, the output of minor publishers, and various publications in specialized fields.

In serving as the national bibliographic center, the Library expects to provide the following services and products:

Authority Information

1. The Library will disseminate name authority records containing not only the established form of headings and its associated see and see-also references, but also the citations of sources and the information used to determine these forms. Some of this information is now provided by the Library's book catalogs and by the new publication, Library of Congress Name Headings With References, but the coverage is not comprehensive and the data are incomplete even for the headings given. In 1976 the Library will begin to put into machine-readable form complete authority records for all name headings used in the current MARC records along with all new and changed records for nonMARC headings. The authority records will be used to produce an enlarged version of Name Headings in book and microform. In due course, these records will be available on-line through the MARC Distribution Service. Gradually, authority records will be provided for all name headings in the retrospective MARC data base. By 1980, the MARC names file should contain about one million headings.

2. The Library will expand the coverage of Library of Congress Subject Headings to include categories of headings previously excluded from this publication. The 8th edition, prepared according to the older guidelines, has been issued on microfiche and in book form. The data will also be available in machine-readable form in 1976. That list will be kept up-to-date by regular supplements which will be cumulated frequently. In the face of present priorities and staff commitments, the Library feels that it cannot undertake a comprehensive study that would pave the way for a major restructuring of the subject heading system. Such a study might be conducted outside the Library if a highly qualified specialist were available. (I made that remark on Friday in New York to a joint meeting of RTSD and ISAD. Monday morning I had an offer from a specialist to visit the Library to consider working with us in the study of such a major restructuring). In the meantime, the list will continue to evolve dynamically in a way that we expect to be responsive to contemporary needs.

3. The Library will continue to apply LC classification and Dewey decimal classification numbers to all MARC records. This will entail a substantial increase in the output of decimal numbers as MARC encompasses more and more languages and forms of material. It seems unlikely, however, that the Library will provide UDC numbers or that it will undertake to develop a new system of classification.

Bibliographic Data in MARC Form

The Library plans to continue the expansion of MARC's coverage so that all of its current cataloging is put immediately into machine-readable form. At present the input annually is as follows:

Books, English: 78,000 records.

Books, French: 11,000 records.

Books, German: 17,600 records.

Books, Portuguese: 3,000 records.

Books, Spanish: 8,500 records.

Films: 8,500 records.

Maps: 4,000 records.

Serials: 10,000 records.

Current total input: 140,600 records.

According to our present schedule for expansion, we will add in this Fiscal Year '76:

Dutch and Scandinavian language books: 10,600 records.

Italian: 6,800 records.

Romanian: 2,300 records

In 1977 our request to Congress includes:

Books in other roman alphabets: 9,000 records.

Sound recordings: 3,500 records

Music: 2,800 records.

Our '78 request will include the Cyrillic alphabet: 24,000; and in '79, other nonroman alphabets: 39,000, for a total in 1979 of about 230,000 titles. We expect to have all of our current cataloging in machine-readable form by 1979.

Inclusion of records for nonroman alphabet languages assumes a satisfactory resolution of the treatment of the great diversity of scripts and characters. We presently have an LC working group engaged in studying this problem. Beginning in 1976, the Library will also make available the current bibliographic data in its automated Process Information File.

Dissemination of Bibliographic Data

1. The Library will provide bibliographic data and authority information in a wide variety of forms -- printed, microform, machine-readable -- as needed to meet the requirements of libraries of all types and sizes.

2. While continuing to provide these services to individual libraries, the Library will promote the development of regional networks so that, whenever possible, they can take on the role of secondary distributors of LC bibliographic data on-line. In assisting network development, the Library will encourage building on present systems in preference to the creation of new ones.

3. The Library will continue to take the initiative in providing packaged data such as the book forms of the National Union Catalog, Films and Other Materials for Projection, Chinese Cooperative Catalog, and Monographic Series. Recognizing its responsibility for insuring the continuity and integrity of such services at a reasonable price, the Library will consider relinquishing them only when there is strong assurance that their transfer would not adversely affect the library community. That is a stronger statement than appeared earlier, and it is changed to reflect many comments that I received on that point. At the same time, the Library acknowldeges its obligation to cooperate with major abstracting

and indexing services to build a comprehensive national bibliographic data base.

4. The Library will continue to provide publishers with Cataloging in Publication information so that the essential cataloging data on American publications will be available in the books themselves. The CIP program also has the great advantage of allowing this information to be included in the MARC data base at the earliest possible time.

A Look into the Future

By 1980 when all of the Library's current cataloging is in MARC form, users will consult it primarily through online terminals. The book or microform catalog in a register/index format issued primarily for your use will serve the Library as a system backup device. It should also help to minimize the use of online terminals for certain kinds of routine searches. In addition, the Library may find it desirable to maintain some special-purpose catalogs in card form. The transition to this new system will be evolutionary, not revolutionary.

When the automated system has a proven capability to meet the Library's needs, new cards will no longer be added to the Main and Official Catalogs. Of course, these catalogs will remain indispensable guides to LC holdings not represented by MARC records. Eventually, the better of the catalogs will be published in book or microform after errors in filing arrangement have been corrected. Meanwhile, the MARC data base will be considered to be complete for cataloging purposes and new entries will be tailored only to its requirements.

By treating the MARC data base as self-contained, the Library will be free to undertake many desirable projects to enhance it qualities: alteration of older name headings that are incompatible with the current rules, adoption of international standards for romanization, and more vigorous improvement of subject headings. But even more important, reliance upon the MARC data base will enable the Library to make complete name and subject authority information readily available. This, in turn, will promote truly efficient decentralized input to the national bibliographic data base.

Naturally, changes stemming from these actions will affect all users of LC cataloging data. Over the years the Library has routinely made changes in its name and subject headings and its classification numbers, and libraries have generally been able to accomodate to them. Although the volume of changes will be significantly greater when we begin to rely upon the MARC data base, their effects should not be severe. Whatever the immediate difficulties may be, they will be more than compensated for by the long-term benefits of bibliographic control and the vastly increased potential for collaborative effort on a national scale.

* * * *

MS. AVRAM: Automation in the Processing Department depends on the close cooperation of many units (Figure 1). The most active are the following:

1) the MARC Development Office is responsible for automation in the Processing Department from the point of acquisitions through the cataloging continuum up to the point of product distribution;

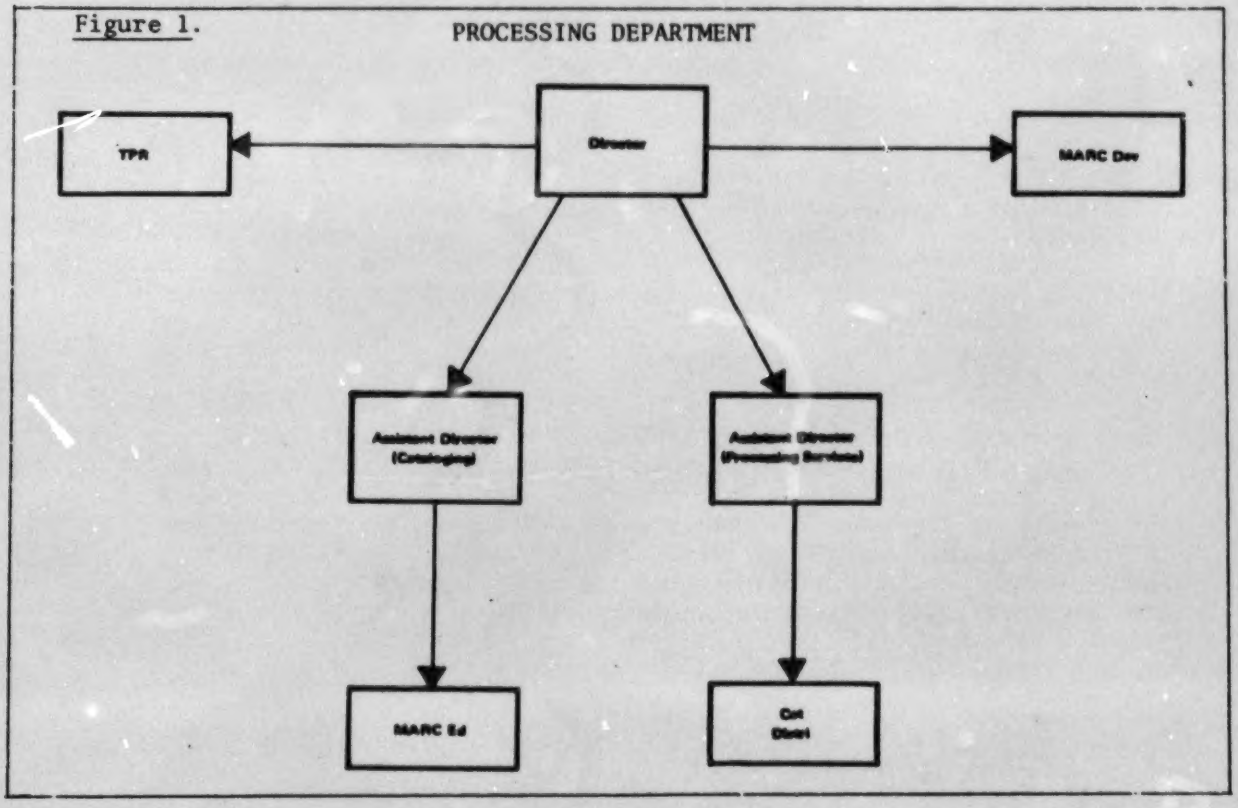
2) the Technical Processes Research Office conducts research for the Processing Department in general but, in addition, serves as the right arm of the MARC Development Office in the research needed to automate complex bibliographic operations;

3) our Serial Record Division, which is not shown on the chart has the responsibility for the assignment of the ISSN and key title for all serial records with U.S. imprints;

4) the Cataloging Distribution Service Division is responsible for the distribution of all cataloging products--photo-composed book catalogs, machine-readable data, printed catalog cards to units of the Library of Congress and also to the library community; and, of course

5) the MARC Editorial Division is responsible for the creation of the records in machine-readable form..

Figure 1.



The MARC Development Office has a dual responsibility. We are very much involved with the automation of a very complex set of technical processing functions to put information under bibliographical control and make this information available to the staff of the Library of Congress, to Congress and to the users of the Library's collections. But in addition, and equally as important, we have a national responsibility to make all this data available to the nation's libraries.

Because the operation at the Library of Congress is so complicated, we have approached automation as a three-pronged effort (Figure 2). We have, and are continuing to automate certain functions to assist units of the Library in the short term. We are able to produce on-demand bibliographic listings; book catalogs for the Main Reading Room, the Science Reading Room, and the Motion Picture Section; the index to the Monthly Checklist of State Publications and other products.

Figure 2. PROCESSING DEPARTMENT AUTOMATION EFFORT

A Three-Pronged Approach

1. LC Short-Term

   On-demand bibliographic listings
   Book catalogs
   Main Reading
   Science Reading
   Motion Picture Section
   Index to the Monthly Checklist of State Publications

2. LC Long Term

   Core Bibliographic System

3. National and International Library Community

   MARC Tapes
   Printed Products
   On-demand Services -- Online and Offline
   Development of Standards

Our long-term plan, which Lucia will describe in greater detail, is our Core Bibliographic System. Our MARC tape service, printed products, on-demand services, online and offline, and, of course, our great activity in the establishment and development of standards, are indications of our national and international commitments.

I would like to define for you our two major efforts--the Core Bibliographic System and the National Bibliographic Service. Our Core Bibliographic System (Figure 3) is a system to control bibliographic information. Its advantages will be: 1) more expeditious production of bibliographic information; 2) bibliographic information available through a greatly expanded variety of access points; and 3) bibliographic products available in a variety of forms--printed, microform, machine-readable form, and online. We are building this system by implementing individual modules which are immediately useful and which eventually will be integrated into the total system.

Figure 3.	Core Bibliographic System (CBS)
Definition:	System to control bibliographic information.
Advantages:	More expeditious production of bibliographic information.
	Bibliographic information available through a greatly expanded variety of access points.
	Bibliographic products available in a variety of forms -- printed, microform, machine-readable form, online.
Method:	Built by implementing individual functions designed to be immediately useful which are then integrated into the long-range system.

The National Bibliographic Services (Figure 4) is a service to support the nation's libraries. Its advantages are: 1) the decrease in costly duplication of bibliographic processing, 2) the provision of tools for cataloging so that cataloging can be performed in a more consistent manner, and 3) the provision of a national interlibrary loan tool.

Figure 4.	National Bibliographic Service (NBS)
Definition:	Service to support the nation's libraries.
Advantages:	Decrease costly duplication of bibliographic processing.
	Provide tools for cataloging in a consistant manner.
	Provide a national interlibrary loan tool.
Method:	Use of data from external sources.
	Make all data from the CBS and external sources available in a variety of forms.

We will build this system using data from external sources, i.e., data from the international bibliographic community, as well as from our own national community, and we will make all the data from the LC Core Bibliographic System and the external sources available in a variety of forms.

* * * *

THE CORE BIBLIOGRAPHIC SYSTEM

Lucia Rather
Assistant Chief, MARC Development Office

When you talk about the Core Bibliographic System you are really talking about the internal processing at the Library of Congress. I think this is a truism, but I am going to repeat it because sometimes I think we forget it. The quality and the quantity of LC services to the national libraries, that is, libraries outside the Library of Congress, are only as good as the products produced by the internal system. In other words, our national bibliographic service is based in large part on the soundness and on the quality of our Core Bibliographic System. That is why we are emphasizing that to you today. If we begin to slip up internally, then we will slip up externally.

Figure 5. MARC DISTRIBUTION SERVICE
Present Coverage - 1975

Form of Materials	Dates Covered	Approximate Volume per Year	Accumulated Totals per Year
Books, English	1968 -	78,300	
Books, French	1973 -	11,000	
Books, German	1975 -	17,600	
Books, Spanish	1975 -	8,500	
Books Portugese	1975 -	3,000	
Films*	1972 -	8,500	
Maps*	1973 -	4,000	
Serials*	1973 -	10,000	
*All languages			140,900

Figures 5 displays some of the MARC statistics that Mr. Welsh referred to earlier, indicating the present coverage of our MARC tape service. The MARC tape service was the very first real bibliographic automation activity at the Library of Congress. As you probably all know, we began with a pilot in 1966 and went on to a full-scale distribution service in 1969. In 1969 we began inputting all of our current English language records. At the time that seemed to be a very small amount, and most people at the Library of Congress did not really feel that this data base was of much use. Here we had 65 years of cataloging, and what use was one year of English language materials cataloged?

But time has gone by, the future is longer than the past, and we now have about seven years of data in machine-readable form; we have expanded to the languages that you see in Figure 5. And we discover from a number of our users, both inside and outside the Library, that we are probably indeed covering or have covered a large portion of the records that are being used in the United States today.

In addition to the English, French, German, Spanish and Portugese books that we are covering at the present time, other forms of material--films, maps and serials going back to 1972 and 1973 are being input. These tapes created at the Library of Congress are, of course, available through our subscription services.

Figure 6. MARC DISTRIBUTION SERVICE
Projected Expansion

Form of Materials	Dates Covered	Approximate Volume per Year	Accumulated Totals per Year
Books, Dutch/Scan.	1976 -	10,600	
Books, Italian	1976 -	6,800	
Books, Roumanian	1976 -	2,300	
			160,600
Books, other Roman	1977 -	9,000	
Sound Recordings	1977 -	3,500	
Music	1977 -	2,800	
			175,900
Cyrillic Alphabet	1978 -	24,000	
			199,900
Other Non-Roman Alphabets	1979 -	35,000	
			234,900

Figure 6 summarizes our projected expansion, showing that we hope to have most, if not all, of our current cataloging input by 1976.

Figure 7 a very brief summary of our MARC subscribers today. As you can see, we have 65 subscribers to the book service, 15 to films, 24 to serials, and 14 to maps. I think it is rather interesting that 15 of these subscribers are commercial services and 22 of them are outside the United States.

Total number of primary subscribers is not large. However, we estimate - and we have no way of really confirming these figures -- that there are approximately 1500 libraries that are benefiting from the MARC tapes in one

Figure 7. MARC USERS (1975)

Primary Users		Secondary Users
Books	65	Estimated 1500
Films	15	
Serials	24	
Maps	14	

Notes: 15 commercial subscribers
22 outside U.S.

way or another. They are getting tapes from other libraries; they are getting products from tapes from other libraries. Of course, a major example is OCLC, which is acquiring MARC tapes and making the records widely available. Many of the commercial services are doing the same thing.

In order to produce all of these different kinds of tape services, we have had to develop a series of formats. The original one is the MARC books format. Other formats have been developed since that first format began: one for serials, one for maps, one for films, and one for manuscripts. We do not have a manuscripts distribution service, but there is a manuscript format. These formats are all developed to be directly reactive to the form of material. However, they are all compatible; they can in general be put together to form one generalized format, so that any system wishing to use a generalized system to process all of its material can combine these formats to produce a generalized format and basically use one set of software.

Now I would like to talk a little bit about the use of these MARC tapes at LC. We used to say that LC was the largest non-user of MARC tapes in the United States. This was true for about the first two years after we started with MARC. We expended all of our energy in creating the tapes that were used outside the Library of Congress. But since then the use of the MARC tapes has, you might say, infiltrated the Library of Congress. I will describe some of the ways that we have used this tape data base.

We developed what we called a generalized retrieval package, which has three main components. The first is a program called the MARC Retriever. This is a program which searches a tape file in batch form. It is a very sophisticated little program. You can search on any tag, any indicator, any subfield code, the date in any field, or any combination of any data, tags, and subfield codes. For example, if you wish to, you could query the data base for all books published in 1973 in English, in translation, in the field of literature that are not juvenile. This may seem outlandish, but we have done this for quite a few years for a professor at the University of Indiana who every year compiles a chapter on U.S. literature in translation. If he did not have this service from us, he would spend his time going through bibliographies, proof sheets, and so forth. He can pay us

and we do a one-time run against the data base and simply give him a fairly complete printout combining all of this information.

Figure 8. USES OF MARC RECORDS IN LC

MARC RETRIEVER

CURRENT AWARENESS LISTINGS

1) Mainland China
2) Population
3) Africa
4) Afro-Americans
5) Children's literature
6) Reference books
7) Conferences
8) Titles in translation
9) East Central, Southeastern and Eastern Europe

Figure 8 shows some of the current uses of the MARC retriever in LC.

All of the listings shown on this figure are done on a monthly basis. In other words, we search the most recent month's records. For example, we provide listings for the area specialists in the Library of Congress, such as those for Mainland China, Africa, and Eastern Europe. This is done by searching the geographic area codes in the MARC data base. We feel that the geographic area codes are one of the big advantages that we have gotten from the MARC data base. It is very hard to do geographic area searches in our existing card catalogs. If you think about it, if you wanted everything on Africa and you were using a card catalog, you would have to look under everything under Africa, every country in Africa and every city in Africa, every natural feature you could think of in Africa, every major subject heading subdivided by Africa, or the countries in Africa. It would be an exhaustive and exhausting search. Using our geographic area codes, by which with each record we give a code that tells the geographic area that the book is about, you can simply pass the entire data base and retrieve those things that are on Africa, Mainland China, or Eastern Europe.

Some of the other searches are in the fields of children's literature, conference publications, titles in translation (which is used for the list we submit every year to UNESCO for the Index Translationum). This is merely a sample of the kind of monthly listings we produce.

Figure 9 indicates what we call onetime searches. People have asked us to do them one time; they are not done every month. Some of them are rather interesting; the two on statistics on Africa and the economics of five African countries were done for the General Accounting Office when they were auditing the Peace Corps. The American maps is run against the maps data base.

Figure 9. USES OF MARC RECORDS IN LC-2

MARC RETRIEVER
"once-only" listings

1) Festschriften
2) Statistics on Africa
3) Economics of Five African Countries
4) Directories
5) 1972 CIP Records Not Yet Published
6) Chronological Subject Subdivisions Requiring Changes
7) All American Maps Published in 1971
8) Caribbean Economics and Political Affairs
9) Environmental Economics (and of Those Selected, Which Ones Were Issued by the Environmental Protection Agency)

The Caribbean economics and political affairs is an example of the kind of search that we do on-demand for the Congressional Research Service.

Some of our searches have resulted from serendipity. One year when we went to make the run on all the books in English and in translation that were not juvenile, and when I was given the printout by the programmer, it looked a lot thinner than I expected. When I looked a little more closely, I discovered he had asked for everything that was juvenile. What I had there was a list of children's books in translation. I was about to throw it away when I thought I would call up the Children's Book Section and see if they would be interested. I discovered that they were in the process of compiling a bibliography of children's books in translation published that year because ALA was giving an award for the best children's book in translation. So that got them interested. And ever since then we have been building a children's data base for them, for which they can do various searches for such things as books in various languages or different types of books.

That is one aspect of this generalized program, the Retriever itself. Once a user has asked that a search be made, he can get his output in a variety of different forms. He can ask that it be sorted by main entry, title, call number, or anything he would like to have it sorted by. In addition, he can ask for the product to be printed in a variety of different ways. He can get it printed by the computer on cards; he can get it printed on listings, one column, two columns, three columns. He can also specify whether he wants the full bibliographic record or whether he would just like a short record, perhaps the author, the title, the date of imprint, and the subject headings.

This is especially valuable, because sometimes you may inform a user that you got 1,500 hits and if the full record is to be printed out, it

would cost the user $330, but if the user just wanted a short record, we can do it for about $100. So the user does have the flexibility of requesting less than a full record.

Most of these runs are done for the Library of Congress; they are done for the Reference Department, the Congressional Research Service, and quite often for the Processing Department when we are investigating what the effects of various cataloging changes would be to the card catalog.

We also do these searches on demand, on a cost basis, if people write to the Library of Congress, but that has not been an extensive service. Searching the whole data base is an expensive process, and we do not encourage people to do it. If they want all the books by Mark Twain, we advise them to use the card catalog, but when they do want some of these things that you cannot get from using the card catalog, the Retriever has proved to be a very valuable way of using the data base.

Figure 10.

BOOK CATALOGS

Computer-Printed

Main Reading Room Catalog

11,000 monographs, 2,600 serials
Titles in the reference collection of the
Main Reading Room, Library of Congress

Science Reading Room

4,000 monographs, 750 serials
Titles in the reference collection of the
Science Reading Room, Science and Technology
Division, Library of Congress

Figure 10 shows some of the other uses of the MARC tapes at LC. In the course of the years we have put several of our smaller collections into machine-readable form. It has been done for two reasons. You may find it difficult to believe, but for many years we had no catalog of the Main Reading Room collection. The reference alcove specialists knew what was in the collection and it was really up to them to direct people to the proper place. And so one of the early projects that we carried out was to put the entire Main Reading Room collection into machine-readable form. The Main Reading Room staff gets printouts via the computer, by author, title, call numbers, and so forth. There are about 11,000 monographs and 2,600 serials in the data base. Actually that data base itself, I believe, will be available for sale later on in the year to people who would be interested in what the Library of Congress thinks is valuable to put into its Main Reading Room.

We did the same for the Science Reading Room, where there are 4,000 monographs and 750 serials. And for this collection we also produced book catalogs printed on the line printer. They are not elegant to look at, but they can be very useful. The records are sorted according to a program called LIBSKED. LIBSKED allows us to arrange bibliographic records according to the current filing rules used at the Library of Congress for its computer-produced book catalog.

Figure 11.

BOOK CATALOGS

Photocomposed

Library of Congress Catalog: Motion Pictures and Filmstrips

LC's first attempt in using MARC Data for book catalogs on a production basis.
Estimates 10,000 records per year.
Three quarterly issues, annual cumulation, and quinquennial cumulation.

Figure 11 refers to photocomposed book catalogs which we are producing, and these are the beautiful ones. I should not use this in the plural at the moment, all we are doing is the films catalog; we decided to start small. We wanted to start with a catalog where we were putting everything in one area into machine-readable form. Films is a small data base; we put about 10,000 film records a year into machine-readable form. And so now the films portion of the National Union Catalog is published each quarter via completely photocomposed methods. We take the MARC tapes of films, process them on the Videocomp in the Cataloging Distribution Service Division and produce those catalogs in that way. We produce quarterly issues, annual cumulations, and eventually we will issue a quinquennial cumulation.

Another use of the existing tape system is the printing of the cards. All the currently produced cards for which there are MARC records com through the Cataloging Distribution Service's Videocomp system. This is helpful in a number of ways. It cuts down on the amount of hot type setting that we have to do, and it also allows us to produce in the very same operation the cards with the overprinted headings to file into all the Library's different catalogs.

That is what we have been doing with the tape data base all these years. And we have been getting quite a lot of use out of it, but we have known that the tape data base would not suffice forever. As that data base keeps getting bigger and bigger, searching via the MARC Retriever takes longer and longer and becomes more and more expensive.

We have been working for the last few years on putting the MARC system

online. We have been developing a system that we call the MUMS System. That means Multiple Use MARC System; it is designed to be a generalized system that will be hospitable to various kinds of applications or uses of the automated data base.

The very first of the projects that we began with was to help out in the creation of the MARC records themselves. The MARC records at present are created at the end of the line. After all the cataloging is done, we take the full bibliographic record and put it into machine-readable form. At present that is a tape batch process.

Any of you who are involved in tape batch input know that you input the record and then you get a printout and you proof it, and when you find mistakes, you mark the mistakes and send it back to the typist, who inputs the corrections and it comes back. And you proof it again, because typists frequently make mistakes when they input the corrections. And then those corrections are corrected and then you get the printout again and you proof it. And this is done at the rate of about one cycle a day. When the record is finally correct, you mark it as correct and it goes back to the typist and she keys in the number again with a verification symbol so that it will be moved from the work file to the full MARC data base. This recycling of records gets to be extremely difficult. It was possible to live with in the beginning when we were doing about 60,000 records a year. But now that we are doing 140,000 records a year, the recycling is unbelievable.

The MARC redesign is an effort to help out by the use of online input and update in the creation of MARC records. We are not going to input them initially online. We are going to input them initially using our offline devices, the MT/STs. We will have a typist type what she sees from the catalog card; it will go into the system, run through our format recognition programs, which create a full MARC record and will come to the MARC verifier as a full MARC record, and hopefully correct. But maybe not correct. So the first thing the verifier will have to do is proof it. The verifier will mark on the printout the corrections that need to be made, and then will take the printout to a terminal, call up the records one by one, make the corrections online right there at the terminal, check each record to see if it is correct, and if so, will hit the verify button. The record will be completed right then and there, no more recycling of paper. We are hoping that this will cut the average time of creating a MARC record drastically.

The development of this system has also meant the development of terminals which would display the Library of Congress's full 176-character character set. This character set is really the ALA character set. It was developed jointly by LC and ALA. It allows us to input and display all the major roman-alphabet languages. I do not think it will include Lappish and a few others, but again, it will include the major roman-alphabet languages and also the nonroman-alpahbet languages in romanized form.

We hope to begin limited production of this system later on this fall. We call this the MARC redesign. We have to be very careful when we use this phrase because people think we are redoing the MARC format and everyone gets uptight. But we are not. This is merely a redesign of our internal operation.

The second use of the MARC online system is what we call the MARC search system. This is what you might call a temporary expedient. We are still inputting the records offline, but then we are taking that tape data base and loading it onto disc and indexing it.

At the present moment we have the entire MARC data base online; there are over 600,000 records in that data base. Those of you who came to the Library of Congress for the tour yesterday afternoon saw various people searching that data base and carrying out various duties.

The data base is currently indexed by three different keys. You can search it by the LC card number, you can search it by a three-three author/title key and a three-one-one-one title key. It has a rather interesting feature, I think. If you put in one of those keys, such as "Uni, Rep", which would be the key you would put in for anything that began "United States" anything and started out "Report" on anything, you would immediately probably get two or three thousand hits. And the system is not equipped to cope with two or three thousand hits; in fact, at the moment it will only cope with 30 hits. The first thing you will get back is a message from the computer saying you have too many hits. You can do something about this. You can resubmit your search and you can add as a qualifying term any word or portion of a word that you know of that will appear in any field in that record.

For example, if you were looking for a report of the United States Bureau of Water Pollution (I have no idea if such exists) you could key in "Uni Rep" and then you could key in the main entry equals water pollution. And then the system would go back and it would select only those things beginning "Uni" and "Rep" that had "water pollution" in the main entry. And this, of course, immediately would cut your search down to a handful of items that you could easily scan. We plan in the fairly near future to add keys for personal names, so that you can search all books by a given author, and also for corporate names.

Figure 12 shows where this system is being used in the Library of Congress at present. It is being used in the Bibliography and Reference, and Correspondence Sections of the General Reference and Bibliography Division. It is being used in the Science and Technology Division by the Information Services Specialists, by the Loan Division, and by the Union Catalog portion of General Reference and Bibliography.

All of these are reference areas that generally use either the Public Catalog or the Official Catalog. The new system allows reference specialists who know they want a recent title or know the title is in a language covered

Figure 12.

MARC ONLINE

Reference Department

Bibliography and Reference
Correspondence Section (GR&B)
Information Services Specialist (Sci)
Loan Division
Union Catalog and International Organizations
Reference Section (GR&B)

Congressional Research Service

Library Services Division

by MARC to simply use a convenient terminal in their area rather than going to the Public Catalog. The Congressional Research Service uses it, the Library Services Division uses it quite heavily. Many congressional requests are limited to English language publications and to material that has been published in the last five years.

Figure 13.

MARC ONLINE

Processing Department

Bibliographic Inquiry Unit (CDSD)
Bibliographic Section (Shared Cat)
Cataloging in Publication Program
Decimal Classification Division
MARC Editorial Division
NUC Control Section (Cat Publ)
Preliminary Cataloging Section (Desc Cat)
Process Information Unit (Cat Mgmt)
Technical Processes Research Office

Figure 13 shows where the system is currently being used in the Processing Department. It is being used in the Bibliographic Inquiry Unit of the Cataloging Distribution Service Division. As you know, when you order cards, you do not have to have the card number; you can just give us the author and title. For many years we have been searching mammoth card files for such requests. Now, when the request seems to be for a current title in one of the MARC languages, it is searched against the online MARC data base.

In Shared Cataloging the staff searches reports coming in from outside libraries where people are asking if a book has been ordered in the Shared

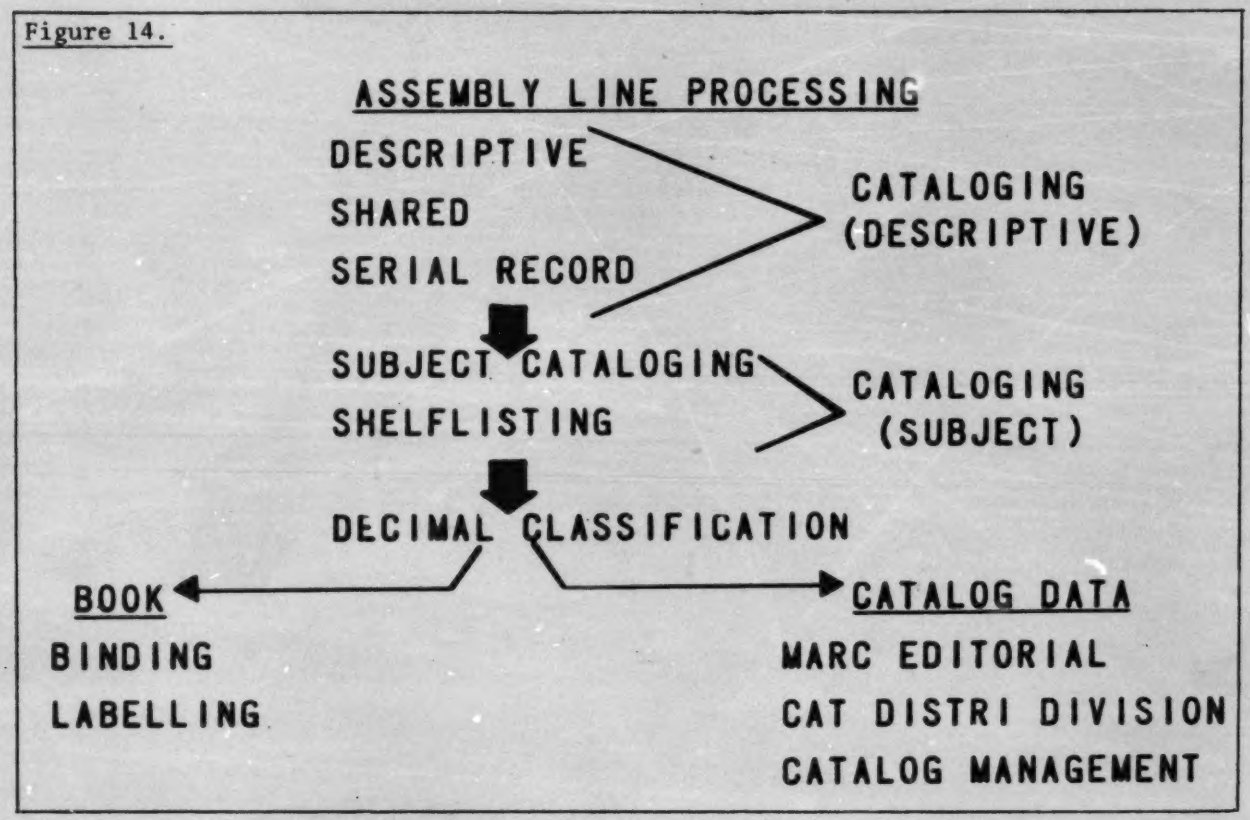
Cataloging System. And if it has, they want the card; and if it has not, they request that the book be ordered. These reports are all searched on the MARC data base.

Catalog Publication is another place where the system is being used quite heavily. This division receives reports for the National Union Catalog, and the staff must first determine if this title has already been cataloged by LC. As the reports are received they are divided into current and noncurrent titles and the current are searched first against the MARC data base.

We have right now 16 terminals that are actually in use more or less most of the time. According to information we get from the Information Systems Office, we are averaging about 6,000 transactions a day on the computer system.

The foregoing has been a picture of the use of the existing MARC data base, the finished bibliographic record after it has been put online. Now I would like to go back a little bit and look at the entire processing of the Processing Department and talk about how we are planning to try to use automation to help us out in our internal activities as we create these records. As indicated in Figure 14, and as most of you probably know, processing at the Library of Congress is an assembly-line process. I do not know how many people get their hands on every book that comes into the Library, but it goes from one station to another, beginning with an acquisitions station and ending up finally in the Cataloging Distribution Service Division.

Figure 14.



Acquisitions is carried out at the Library of Congress in at least three, maybe four, divisions. We get things via copyright, exchange, and gift, gifts from publishers. And, as a last resort, when we cannot get something any other way, we will buy the book; so we have an Order Division as well. The Order Division system has been in the process of being automated for the last few years. The bibliographic portion of that system is complete. In the bibliographic system we input the title just once and get from the system orders that go off to the publishers, follow-up orders when the book does not arrive on time, in process listings to tell us the status of any order, notices to people in the Library who have requested that books be ordered to tell them the book has been ordered, and that sort of thing. We are currently working on the fiscal portion of the system, so that the system will do all of the debiting of the accounts and so forth. That should be complete some time in the spring.

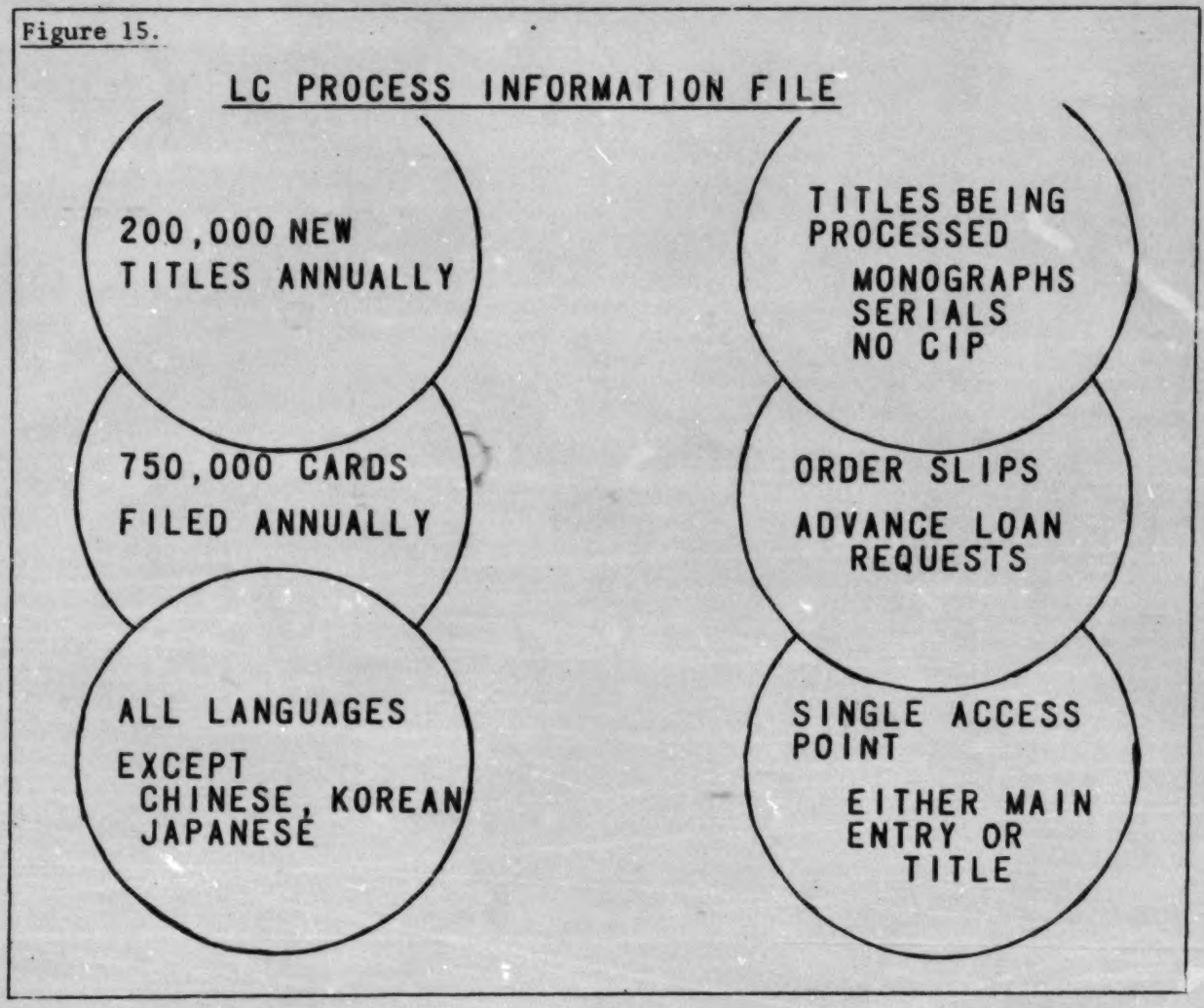
Once the book gets into the Library it has to go through many different processes, and to control this, we have what is known as a Process Information File. As a book goes from one cataloging division to another, it is necessary that we be able to find where the book is. Some of the books that come into the Library are processed very quickly, and some are processed very slowly indeed. We have a priority system. American book trade books and books that are acquired on the Shared Cataloging Program by request of the NPAC libraries get very high priorities and go through the system just as fast as possible. If a congressman wants a book, it really goes through the system in a hurry.

But there are other more ephemeral materials, such as pamphlets from Latin American countries or Balkan countries that we have acquired through our various acquisition systems -- they have very low priorities and they go through very slowly. In fact, some of them do not go through at all. We have some records in our Process Information File that date back to the 1940s for materials that have gotten lodged in backlogs; no one has ever asked for it, and there are always more important things that have to be done. But we have to have a Process Information File to control where this material in process is.

Figure 15 shows what is in that file right now. We add about 200,000 titles to it annually. As each book goes to each cataloger, he sends a card through to the file saying, "I now have the book, bump out the previous cards," so that you can look in the file and say that Mr. John Smith in Subject Cataloging got the book on May 12th, 1972, or whatever. Approximately 750,000 cards are filed in that file each year, in all languages except the oriental languages. We also have in the file order slips and advance loan requests. When we have a congressman who wants a book, we flag the record in the file saying, "Give it a higher priority."

The manual file at this point is single access. The older part of the file is arranged by author; the current part of the file is arranged by title. We use this file for two purposes. When a new book comes to the

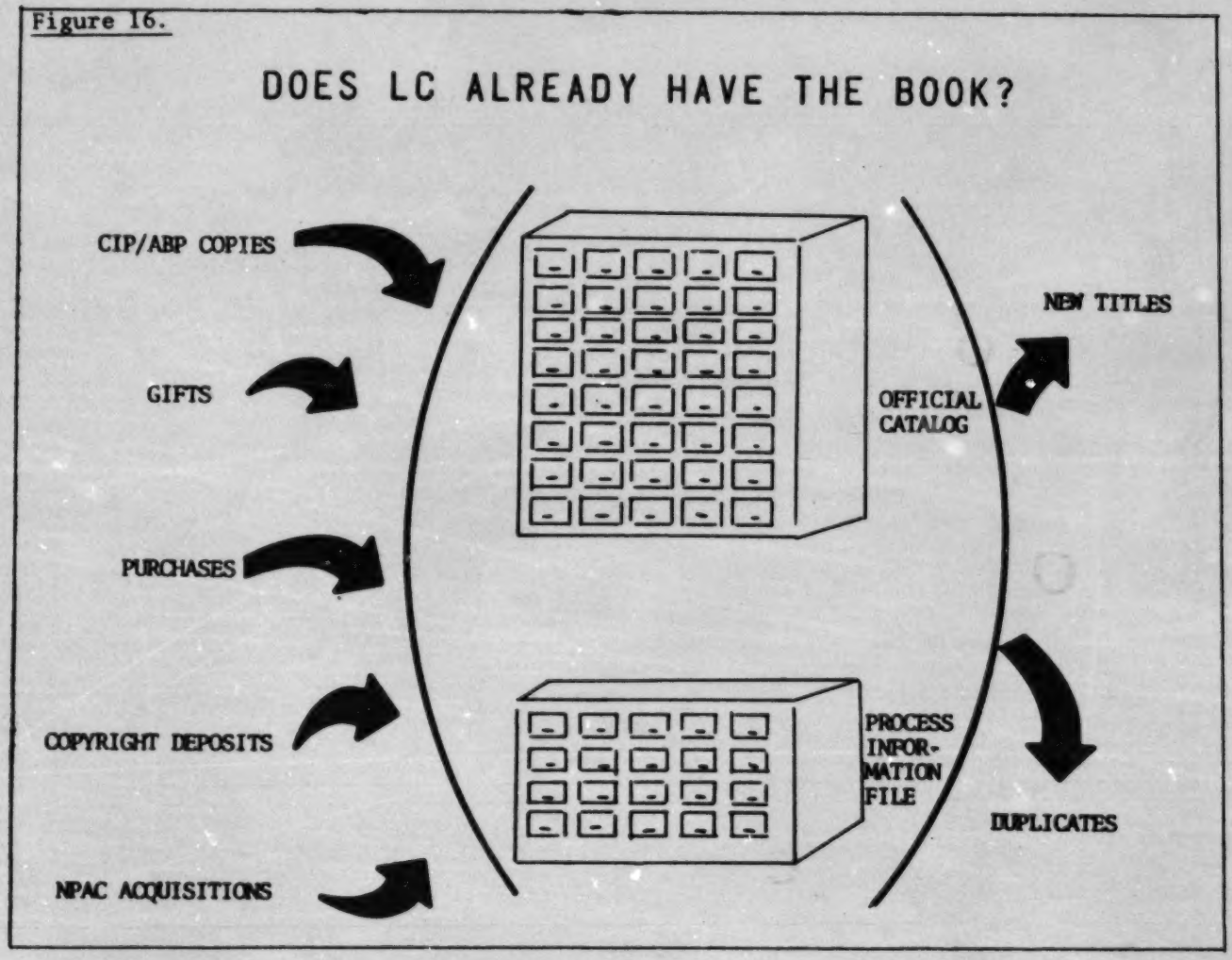
Figure 15.



Library, we want to know if the Library already has the book? Because we acquire things from so many different sources, we acquire a very large percentage of duplicates. So we certainly cannot assume when a book gets to the Library that it is new. We always have to search it first to see if it is in the system. We must search both the Official Catalog and the Process Information File for each new title as it comes into the Library, as shown in Figure 16. The second question is, "where is it?" It may be on the shelf, but it also may be in Mr. John Smith's cataloging arrearage and we want to add this copy to the first one so that both of the books can go through together.

Figure 17 shows in diagrammatic form the progress of a catalog record through the system. Those little boxes at the bottom are what we call our manuscript card, the card the cataloger is cataloging on. Attached to

Figure 16.



each manuscript card is a set of 3x5 cards containing the preliminary cataloging information. As the manuscript card goes through the system, the cataloger extracts a 3x5 card and sends it to the Process Information File. So that is the existing card system.

Once a book has been acquired, it goes to one of our cataloging divisions: Descriptive, Shared, or the Serial Record Division that catalogs serials. It then goes on to Subject Cataloging and Shelflisting, and then to the Decimal Classification Office where Dewey numbers are added. Then the book and the card separate. The book goes for binding and labeling to the shelf, and the catalog card goes now to the MARC Editorial Division where it is put into machine-readable form. Then it goes down to the CDS Division, where the cards are printed, and finally the cards go to Catalog Management where they are arranged for filing in the Official Catalog. So you can see that the book and the catalog record go quite a few places, and you have to keep track of them all the way through.

Figure 17. CURRENT LOCATION REPORTING SYSTEM

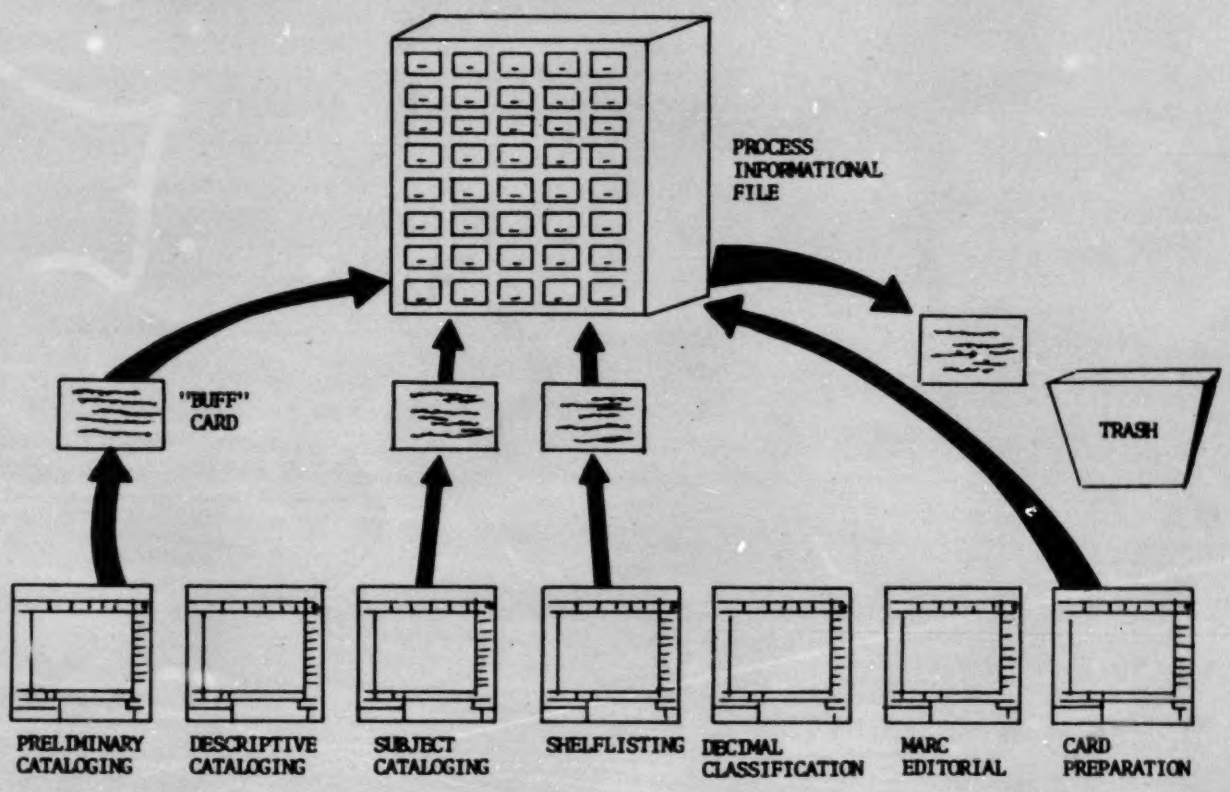


Figure 18 shows some of the problems with the manual file. It is a massive filing effort, and misfilings occur. And once you get one card misfiled, as you probably all know, you are likely to start misfiling behind that card and get two files in the same alphabetic sequence. There is a delay in purging of the files, and sometimes records are not completely purged. This results in records in the Process File for materials which have really completed processing and are in the Official Catalog. In addition, there is only one access point, so you must know fairly accurately the title of the book to find out whether the book is in the Library.

Figure 18. PROBLEMS WITH MANUAL PROCESS INFORMATION FILE

Size of File
- Massive Filing Effort Required
- Misfilings Occur
- Delay in Adding/Replacing Cards
- Incomplete Purging

Single Access Point

Few Location Reports

Single Physical Location

We cannot afford to have all the different stations that the book goes to send a card to the Process Information File; we can only have a selected few. Descriptive, Subject, and Shelflisting are really the only ones that send in a process file card now. In addition, the file is in a single physical location. The Library of Congress is spreading all over Washington; we are now in about six or seven locations. People who wish to find out something have to call up the Process Information File and get someone to search it when they are trying to find out if a book is in process.

We are now in the process of automating this file. Under the new system, when the book has been certified as new to the library, a preliminary cataloger will key into the system an abbreviated record that only includes the author, the title, and the basic descriptive information, but not any of the subject information or the classification number. We will put into this file all those books that are new to the Library of Congress.

Figure 19.

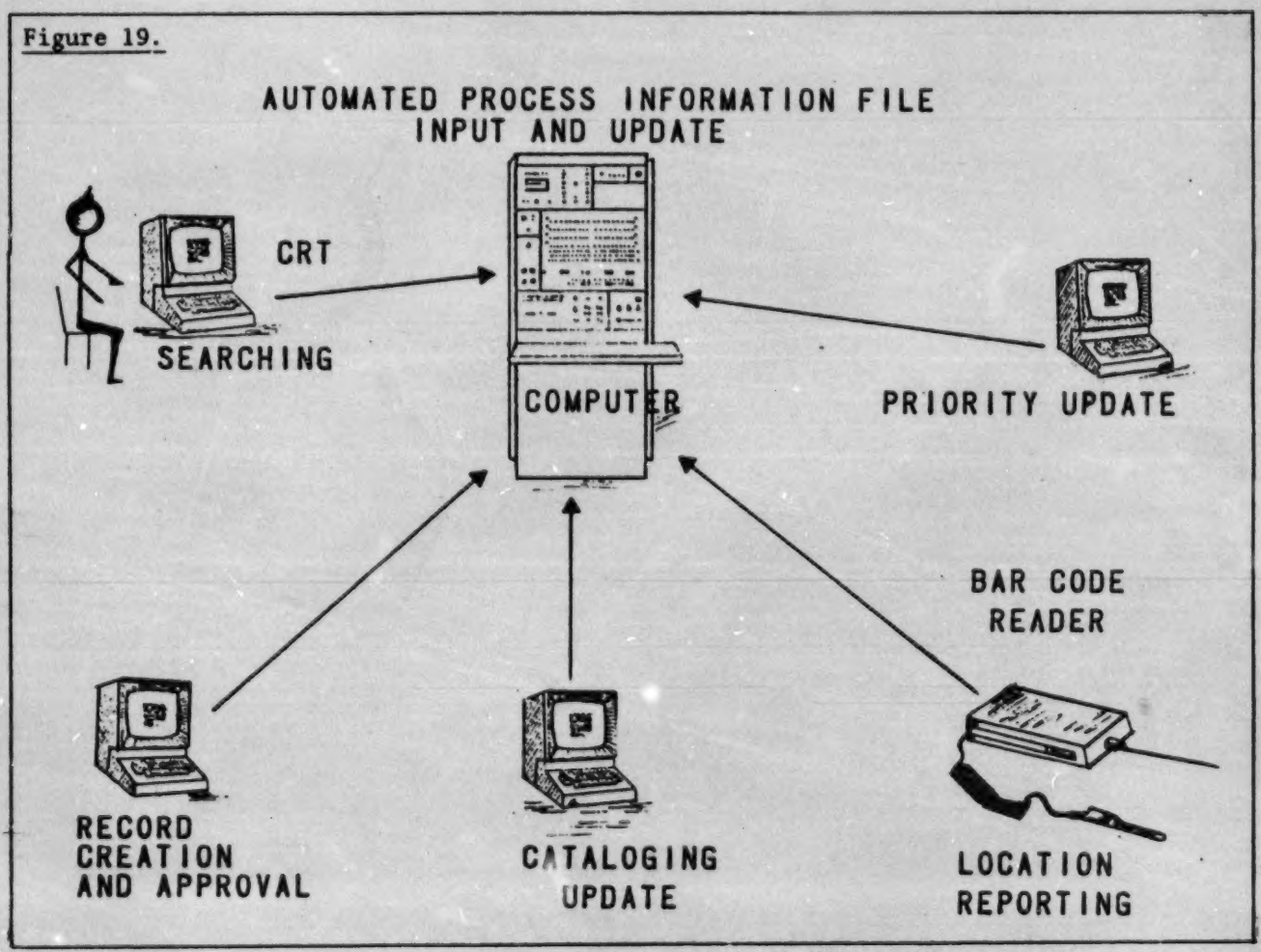


Figure 19 is really a schematic of the current processing system. It shows that we will be inputting via a terminal to the system, and then from the system a manuscript card will be printed out to be used by the cataloger. A bar code label will be pasted on the manuscript card. As the book then goes through the system, each cataloger will be equipped with a badge that will also have a bar code on it. It is the same kine of bar code that you see in supermarkets on products these days. There will be various stations located throughout the Processing Department, and at these stations you will be able to report to the system where the book is at that point.

Figure 20.

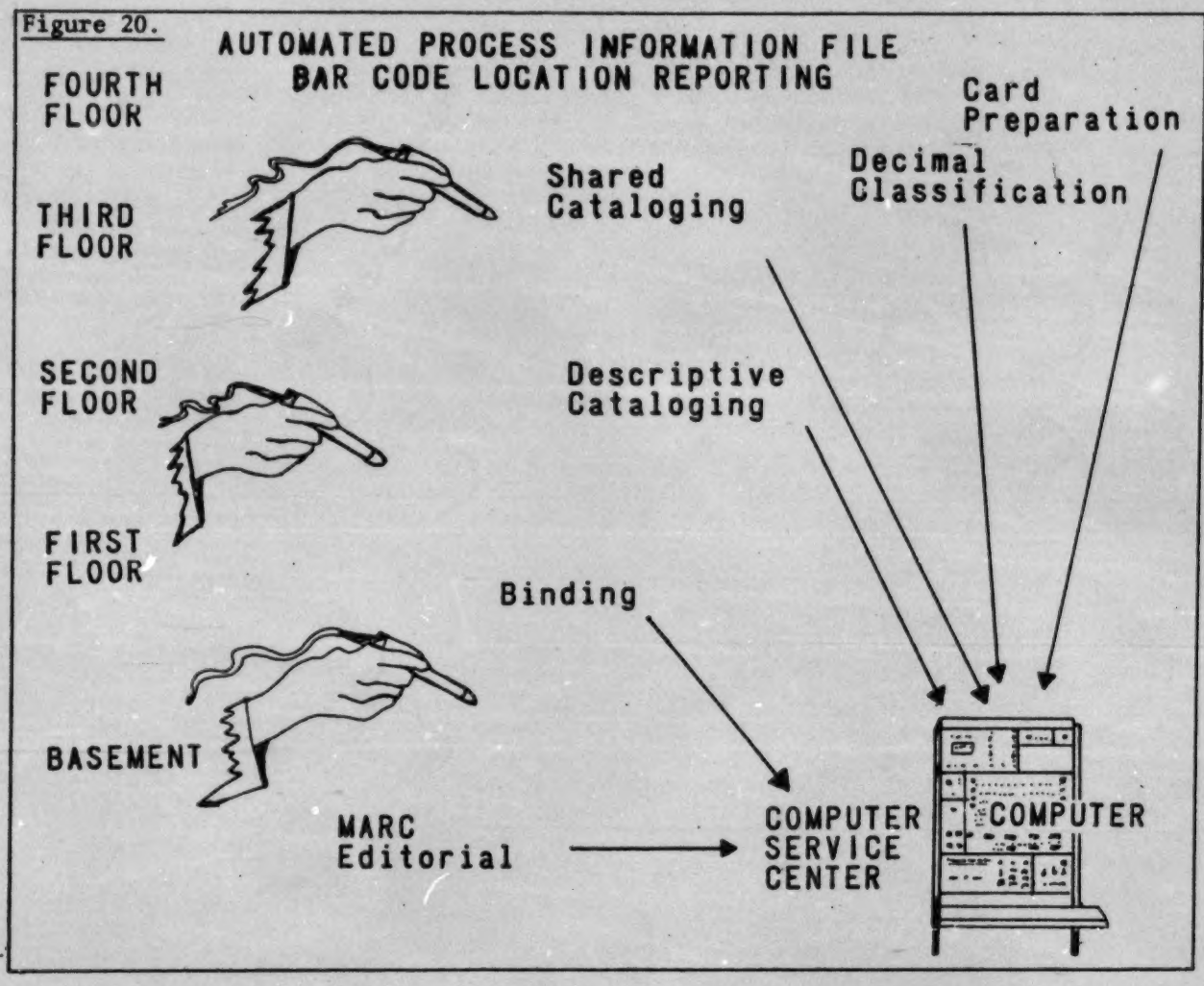


Figure 20 shows how we will report on where the book is in the system. For example, a cataloger gets a book, and the manuscript card and the book will have a bar code label pasted on them. The cataloger will have a badge of his own that will have a bar code on it. He will take it to this station

and pass a light pen over both his bar code and the book's bar code, and the system will say, Mr. John Brown in Subject Cataloging on October 12th has such and such a book. Then if anyone wishes to find that book, they can go to terminals, which will be located all over the Library, call up that record by a variety of search points, and find out 1) Is it in the Library? and 2) Where is it in the Library? thus ending all the filing that we have in the current file.

Figure 21.

APIF SEARCHING ACCESS POINTS

LC Card Number
Author/Title Key
Title Key
Personal Author Key
Corporate Author Key

Series
ISBN

Figure 21 shows the access points to the Automated Process Information File (APIF). They are essentially at the moment the same access points that we have in the MARC search system: LC card number, author/title key, title key, personal author key, and corporate author key. We hope to add before very long series and ISBN.

Figure 22.

ADVANTAGES OF
AUTOMATED PROCESSES INFORMATION FILE

Currency of Process Information File
Manual Filing Saved
Rapid Searching for Many Items
Multiple Access Points
Access from Remote Stations
Several Files Searched Simultaneously
More Complete Location Reporting
Management Information Data Collected
Early Distribution of Cataloging Data Possible
More Efficient Keying of Preliminary Cataloging
Greater Control of LC Card Number Assignment

Figure 22 reviews the advantages of the automated system. For one thing, it should be much more current. There will not be a delay in waiting for somebody to file into that file a card showing that Mr. John Smith has the book.

I will not go through all of the advantages, but the access from remote stations means that someone down in the Map Division or in the Cataloging

Distribution Service Division in the Navy Yard will no longer have to call in to find out; they can use their own terminals to search the file.

And although there will be physically two files, one Process Information File and one MARC File, this will be transparent to the user; he will ask for the record and the system will search both files and tell him where the record is. He will not necessarily have to specify what file he thinks it might be in. So these are, then, basically the advantages of the Process Information File system.

This system has also been dependent upon the development of the terminals with the extended character set. We hope that input to the system will begin in January with English language records, and we will expand as quickly as possible to all the other languages. So you can see that, although we have 16 terminals going now, by January or February we will probably have some 50 or 60 terminals in use at the Library of Congress.

We also plan to make the APIF records available on tape. These records will be available through the MARC Distribution Service. This preliminary information available some weeks, months, or years prior to the creation of the final catalog card, may be useful for acquisitions or other purposes.

Figure 23.

AUTHORITIES - 2

PRODUCTS

- Authority Lists
  - LCSH and supplements
  - Names
  - Microforms
- Distribution service
- Book catalogs
- Miscellaneous
  - Weekly Lists
  - Cross Reference cards

Online

The second major area we are working on right now is in the area of authorities (Figure 23). As you know, we have put our subject data base into machine-readable form. This is the old subject data base that was used to create the Seventh Edition. We had to convert it to MARC form. We have now finally issued the Eighth Edition. I think some of you saw a copy as you went on the LC tour, or maybe you already have your copy. We have also made

it available in microform, and we hope that we will be able to provide updated microforms much more inexpensively and much more frequently than the old printed edition. We have also been producing all the supplements to the MARC subject headings list from that system for the last few years. We also hope to make a tape service available with those subject headings available some time in the spring.

An added advantage to this automated system is that for the first time we have been able to include in that data base what we call the nonprint subject headings. Some of you may think all of our subjects are in that red book, but they are not. We have omitted such things as personal names, proper names, and names of geographic subdivisions, because it simply would have made the list too large, too unwieldy to use. Now that we are able to put the information in microform, we will be able to include this nonprint material and probably make the nonprint headings available in the microform editions even if they are not available in the printed edition.

Another advantage to LC is that we are now printing the cross references to go into the card catalogs from that system. We are developing a similar system for names. This system is being developed first to help LC. We are developing this system to aid the cataloger in the cataloging process. But as a byproduct we hope we will be able to produce the publication Names with References, tapes perhaps with names data, a microform edition of the name authorities, and so forth.

We also plan next to put both the names and the subject authorities systems online. This will mean that our catalogers will be able to search that authority data base online and eventually link that data base with the appropriate bibliographic record. Not only should this system be available for use by catalogers inside LC, but we hope eventually it will be available online outside LC.

Figure 24.

GENERALIZED PUBLICATION SYSTEM

1. Interface with Name and Subject Authority Files to produce cross references

2. Output to: Computer Printer
   Videocomp
   COM

3. Provide Register/Index type catalog

4. Provide Authority Lists

Finally, we want to use this authorities data base in the use of a generalized publication system (Figure 24) which will have several features. It will allow us to take our bibliographic data base, process it

against the authority files whenever we want to produce book catalogs, and generate automatically all the cross references that will be needed for a given issue of that publication. It will be generalized also in the sense that we can use the same set of programs and, by changing parameters, direct the output to a computer printer, to a Videocomp, or to a COM (Computer-Output-in-Microform) device. We hope that it will provide not only the traditional book catalog format that we have had for many years, but also a register/index type catalog. And, finally, we hope to use it in the production of authority lists.

So I would like to conclude with a review of what we are trying to do in this Core Bibliographic System. We want to capture the data at the beginning of the process, and we hope that we will be able to upgrade it and augment it as it goes through the system. We hope that the catalogers will be able to compare bibliographic data against our various authority files and be helped out in the creation of data by automatic verification of various fields against authority files.

Finally, we hope that when we finish the system we will have the production of the bibliographic records by a man/machine mix.

* * * *

THE NATIONAL BIBLIOGRAPHIC SYSTEM

Henriette D. Avram
Chief, MARC Development Office
Library of Congress

Lucia has described the Core Bibliographic System to you. I am going to tell you something about our national and international activities and then attempt to draw together the various projects into a cohesive whole.

Figure 25.

MARC DEVELOPMENT OFFICE
NATIONAL ACTIVITIES

Organization	Committee	Task	Participation
American National Standards Institute	Z39.SC2	Formats*	Chairman Member
	Z39.SC37	Country Codes	Chairman
American Library Association	RTSD/ISAD/RASD	Representation in Machine Readable Form of Bibliographic Information	Liaison
	RTSD/CCRC	Rules for Machine Readable Data Files	Member
Music Library Association	MLA/MARC	Music Format	Member
Council on Library Resources	CEMBI	Sharing Bibliographic Information (Books)	Member
	CONSER	Conversion of Serial Data Bases	Member
Advisory Group on National Bibliographic Control	Working Party	Formats for Journal Articles and Technical Reports	Member

*Adopted standard

The staff of the MARC Development Office has been very active in national activities (Figure 25). We know the importance of standards, as do the rest of you, and have given a lot of our effort to this activity. We are working with the American National Standards Institute on the subcommittee responsible for formats and country codes. We work with the American Library Association, principally with the MARBI Committee, which is an interdivisional committee. Other ALA responsibilities include membership on the RTSD Subcommittee on Rules for Cataloging Machine-Readable Data Files, various activities in ISAD, and, work with the Music Library Association on the development of the MARC music format. We were on the committee sponsored by the Council on Library Resources that was concerned with the problems of sharing bibliographic data already in machine-readable form. The work of this committee led to the COMARC project, which I will describe in a little more detail later. We also have set up by the Advisory Group on National Bibliographic Control for the format design for journal articles and technical reports.

Figure 26.

MARC DEVELOPMENT OFFICE
INTERNATIONAL ACTIVITIES

Organization	Committee	Task	Participation
IFLA	Working Group	Content Designators	Chairman Member
ISO	TC46/SC4/WG 1	Character Sets	Chairman Member
	TC46/SC4/W3 3	Filing	Member
	TC46/SC4/WG 4	Bibliographic Formats	Member

Following the development of MARC at the Library of Congress, MARC projects were implemented all over the world (Figure 26). Though the structure of the format used was identical across international projects, the data content of the format was not standard. That is not surprising since there are no internationally accepted standard cataloging rules, or any international subject heading system or classification scheme. In addition, in the international environment, we are concerned not only with libraries but also with national bibliographies which perform a different function than a library. This difference is reflected in the data content of the record.

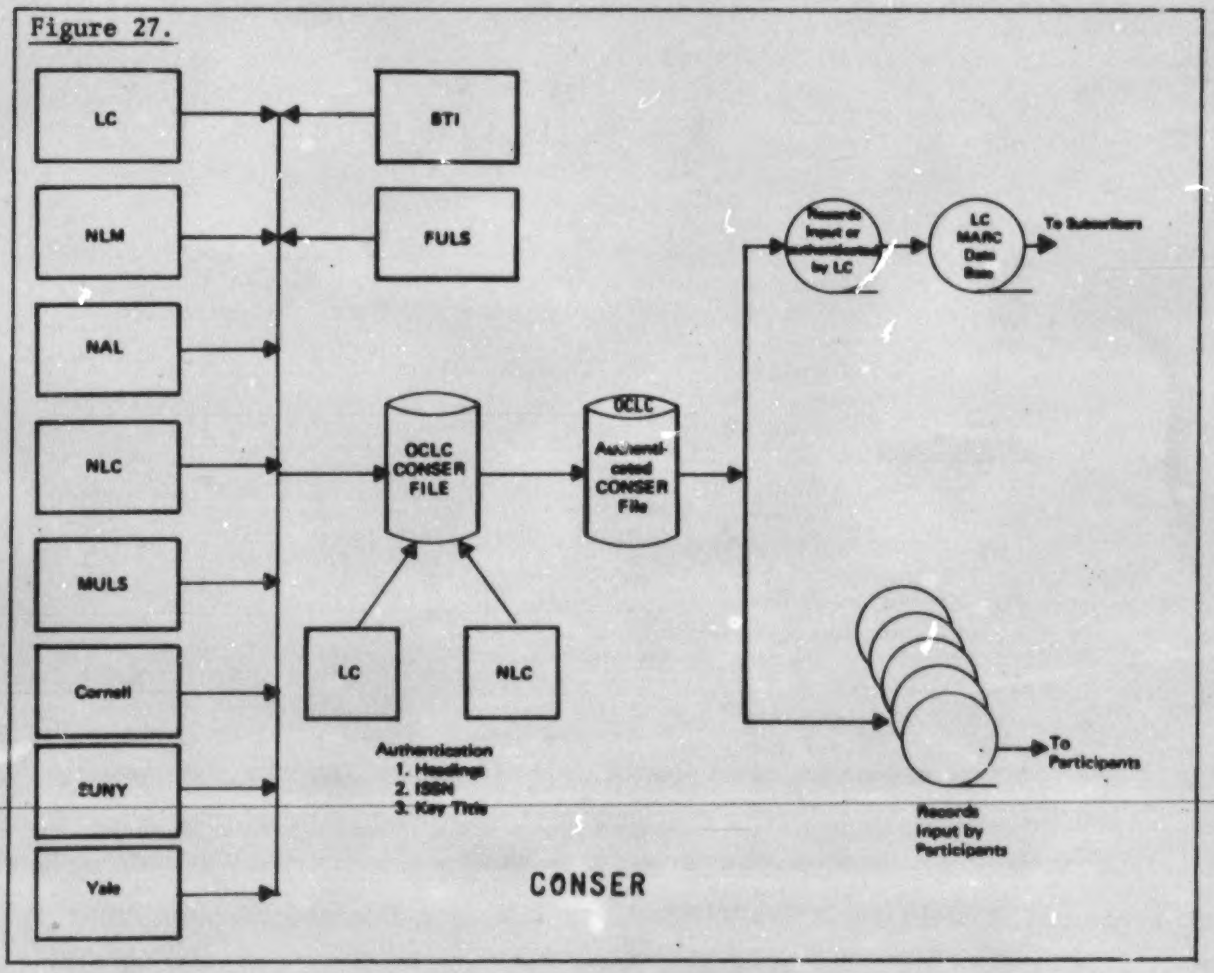
A great deal of MARC Development Office effort has gone into working in the international environment toward standardization. We are part of

a very important working group in IFLA concerned with designing an international MARC format for exchange of bibliographic information across national boundaries. Great strides have been made in this area, and the success to date can be attributed to the ISBD for monographs, which forms the base record for the international MARC format.

We are also working closely with ISO on character sets and file arrangement. We are principally concerned with character sets because of the great amount of work done in this country, and we would like to see the international character set follow the ALA character set as closely as possible.

International work is very meaningful for LC because of our Shared Cataloging program. We see the day when the machine-readable records received at the Library will be input into our system; the bibliographic description will be used as given in machine-readable form and the record will be modified to add LC names, subject headings, Dewey numbers, etc. This will avoid duplicate effort and should be cost beneficial to both LC and to this nation's libraries.

Figure 27.



One of the most significant national projects going on at this time is CONSER; CONSER standing for CONversion of SERials (Figure 27). CONSER is a cooperative effort. Its aim is to build a national serial data base. It is managed, at the present time, by the Council on Library Resources, with the input facility at the Ohio College Library Center. The Library of Congress, the National Library of Canada, the National Library of Medicine, and the National Agricultural Library are all members, as well as several other large research libraries in this country.

The idea behind the project is the input of bibliographic records for serials from all these institutions. The records will be authenticated by the Library of Congress and the National Library of Canada; the National Library of Canada will be responsible for the authentication of Canadian imprints and the Library of Congress for the authentication of all other imprints. In addition, the ISSN and the key title will be added to the bibliographic records by the Library of Congress or the National Library of Canada. Where ISSN and key titles are already in the record, both these centers will authenticate the ISSN and key title. The records will be housed in the OCLC data base. The authenticated record (i.e., those input or updated by the Library of Congress or the National Library of Canada) will be returned to the Library of Congress and will be made available to subscribers through the MARC Distribution Service. In addition, all CONSER records, including those that have not been authenticated, will be made available by the Library of Congress as a separate distribution service.

To satisfy the requirements of CONSER, and in cooperation with the National Library of Canada, the Council on Library Resources, the abstracting and indexing services, the International Serials Data System (ISDS), and other organizations in this country, the MARC serials format was modified. LC will continue to maintain the MARC serials format and publish addenda as needed. In addition, the Library has provided CONSER training, published a CONSER editing guide, and prepared a terminal operating guide for publication by the Council on Library Resources.

Figure 28. CONSER

Two-Year Period - 200,000 Records

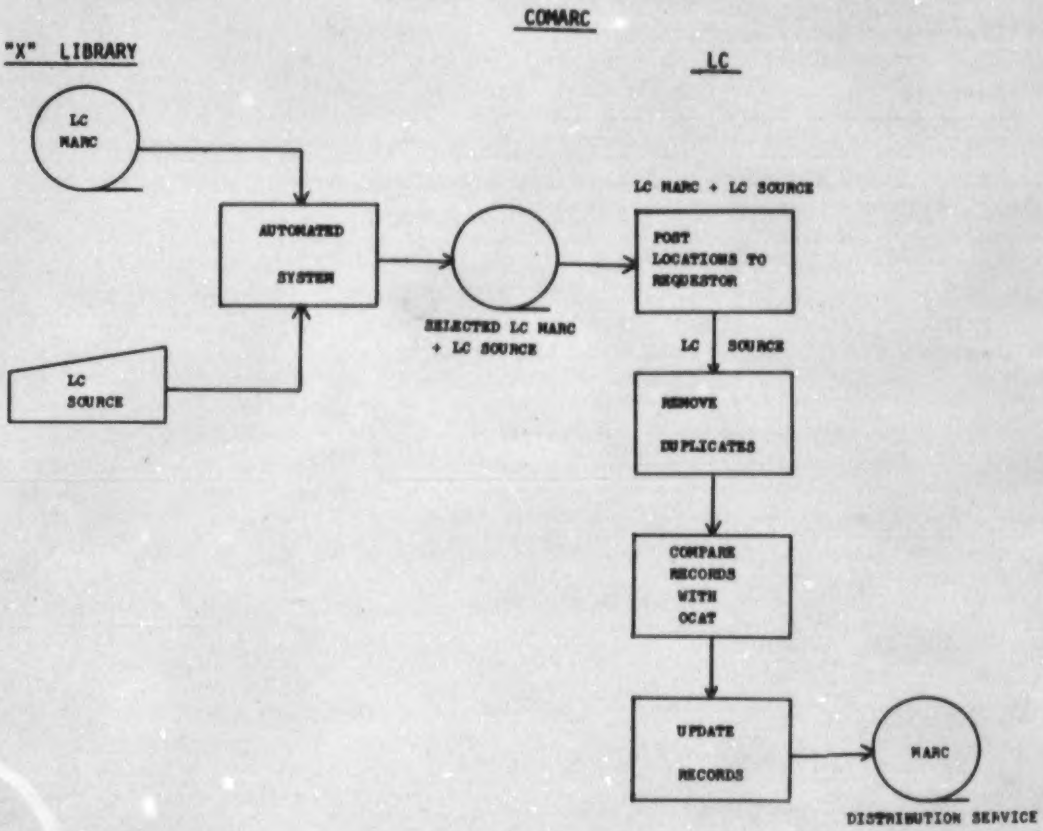
LC New Records	24,000
MULS	81,000
Other	+95,000

As shown in Figure 28, CONSER is a two-year effort. In the two years we hope to add to the data base the Library of Congress MARC serial records, i.e., new serials cataloged and converted to machine-readable form by LC, totalling about 24,000 records. We will be loading the Minnesota Union List of Serials (I believe that this has already been partially accomplished), consisting of about 80,000 records. It is estimated that the other partic-

ipants will input approximately 95,000 records. This will mean that in the two-year period the data base will be somewhere in the vicinity of 200,000 records.

It is planned that this project will be returned to the Library of Congress some time in late 1977. We are in the early phases of exploring the requirements to bring the project back to LC.

Figure 29.



COMARC is another national project (Figure 29). As I noted before, it is the outgrowth of the meetings sponsored by the Council on Library Resources. When one talks about the sharing of machine-readable records among organizations, many problems become evident, and the more you explore, the more evident the complexity becomes. LC proposed to the Council on Library Resources a pilot project wherein the various institutions that are converting LC source data into machine-readable form send the records back to the Library of Congress where the records will be compared with the Official Catalog, updated when required, and redistributed to the library community through the MARC Distribution Service. As MARC expands, COMARC should diminish, since COMARC

covers records cataloged by the Library of Congress, but not included in MARC because 1) the records were either outside the scope of the MARC languages at that time or, 2) the records represent cataloging data from a period before MARC began. Funds were granted and COMARC is underway.

In addition to LC cataloging data (which we call LC source records) converted by another organization, LC MARC records, selected by that organization from the LC MARC tapes for inclusion in its automated system, will be sent back to the Library of Congress. The LC MARC records will be used to post locations to LC's automated Register of Additional Locations. LC source records will also be posted to the Register and then processed to remove duplicate records compared with the Official Catalog and, where necessary, the access points will be updated. The updated records will then be made available at no cost to the pilot participants in COMARC and also through the regular MARC Distribution Service to any organization that is interested.

Figure 30. COMARC

LC Source Records

Criteria for Participation

1) Full Bibliographic Content
2) MARC Format
3) Agreement for LC to Distribute at no charge to LC

Distribution

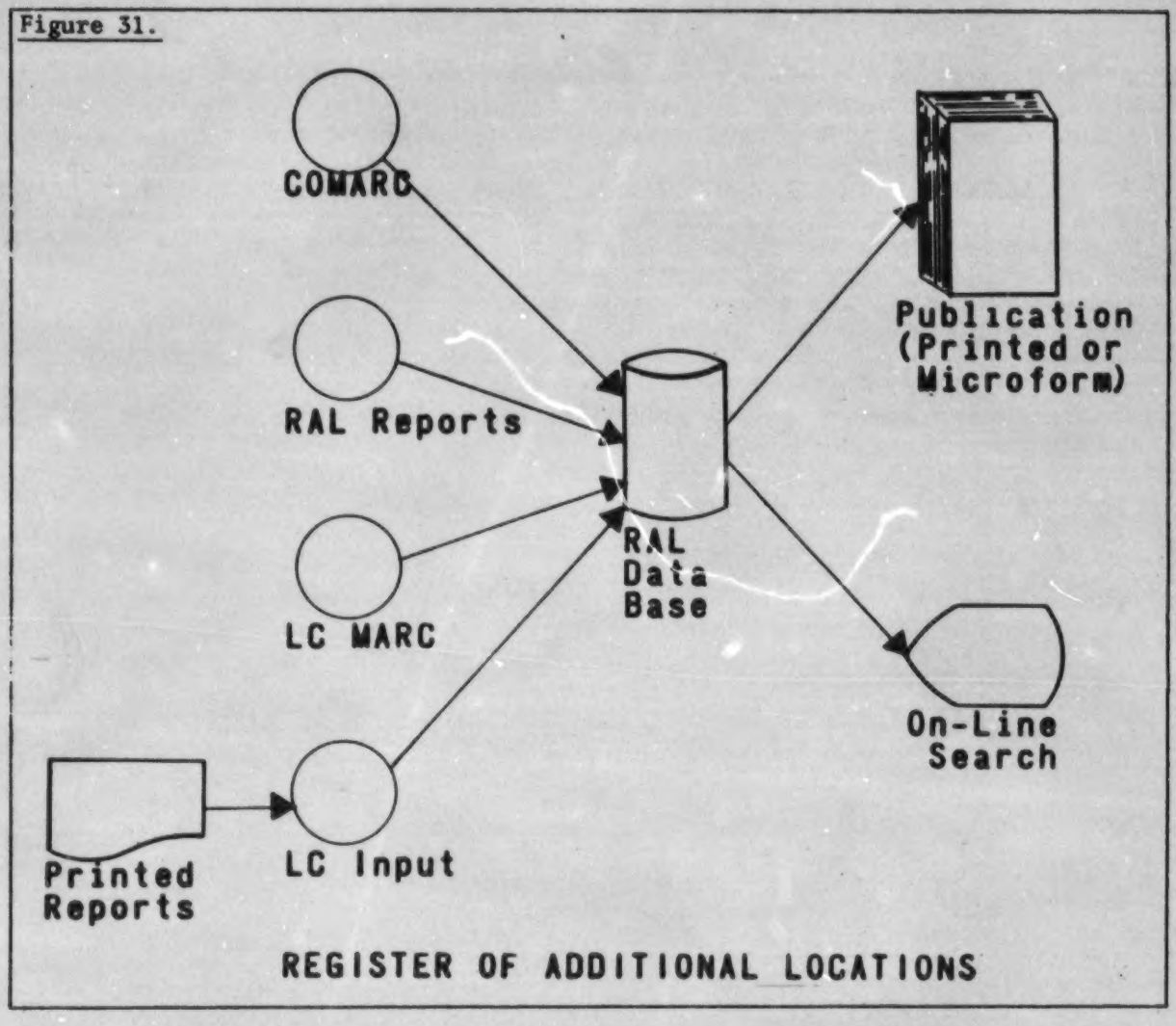
1) At No Cost to Participants during Pilot

2) Also available by subscription to any interested organization

Technical specifications have now been written and distributed to those organizations that have expressed interest in COMARC. Each organization must agree to provide COMARC records to the Library of Congress free of charge (Figure 30) The organizations now participating in COMARC are the Washington State Library, Northwestern University, and Information Dynamics Corporation, and together they will contribute approximately 16,000 COMARC records per year. Yale has also just become a member. Since COMARC is a pilot project, and there was no accurate way to project volume of records available to the project, staff and selected independent of volume. The COMARC cutoff point will be that point where the number of records exceeds the number that can be handled by the number of staff funded by the Council.

The Library has automated the Register of Additional Locations (Figure 31). We now have a data base consisting of titles and locations for titles reported during the period 1968 through 1974, and we are in the process of adding 1975 reports. This voluminous file contains approximately 1,400,000 titles with

Figure 31.



an average of 9.5 locations per title. Some 2,000,000 location reports are added each year.

When the Register system was begun, LC was keying the printed reports received from reporting libraries. The MARC Development Office proposed a pilot project with the New York Public Library wherein New York Public Library would submit to the Library location reports in machine-readable form. This project was initiated, and New York Public Library reports in machine-readable form are now input directly to the Register data base. We hope to expand this to other interested organizations.

The Register file has been made available in printed form. We are now involved with the procedures to make this file available in microform some time in the spring of 1976.

The Register file is used a great deal by the Union Catalog Reference Unit in the Library of Congress, and consequently that file is being made available online in conjunction with the MARC bibliographic file. When a request is made to LC for location information on a title and the LC card number is not known, the MARC file will be searched for the LC card number by author/title or title search key. The LC card number will then be used to enter the Register file for the location information.

Under the COMARC project, LC received funds from the Council on Library Resources to conduct a study to define a format for reporting bibliographic records in machine-readable form to the National Union Catalog. This study was undertaken by David Weisbrod of Yale University. The format is designed for institutions to report titles not cataloged by LC to the national union data base. The reporting format represents full bibliographic content but limited content designation. The format assumes the use of format recognition programs at the Library of Congress to actually tag the records automatically. We hope, under COMARC, to be able to do some experimentation along these lines using LC source records. Much more detailed analysis is still to be done in this area.

Figure 32. INTERNATIONAL MARC

Planned or Operational	Under Study
Australia	Poland
Belgium	Soviet Union
Canada	
Denmark	
France	
Italy	
Japan	
Latin America	
Norway	
South Africa	
Spain	
United Kingdom	

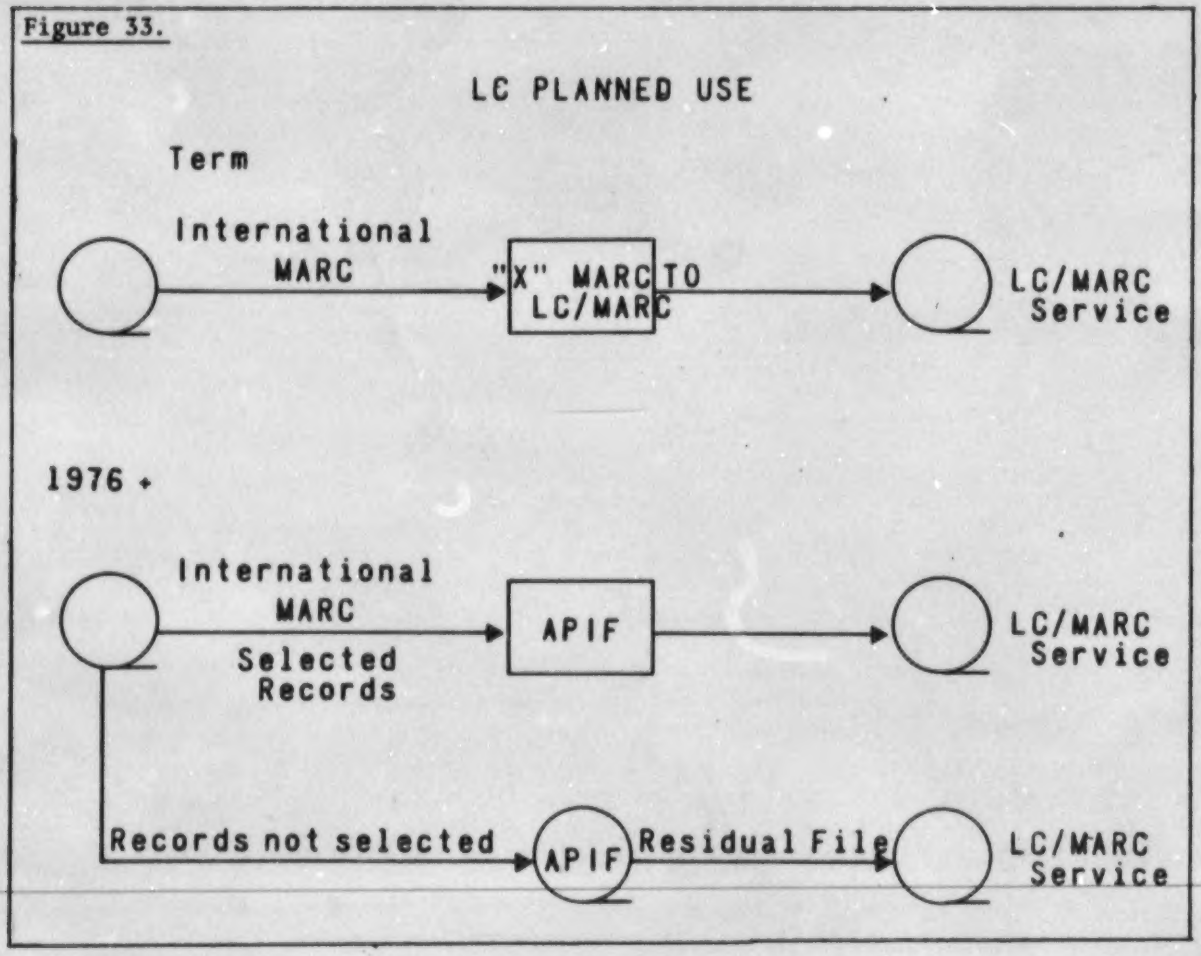
As I indicated before, International MARC is growing very rapidly. We are no longer sure whether the information in Figure 32 is complete. The majority of the MARC projects listed as planned or operational are now operational. There are systems in existence in Australia, Canada, Denmark, France, Italy, Japan, Norway, and the United Kingdom. There may be others also in operation. LC has already entered into an agreement with the National Library of Canada and with the Bibliotheque Nationale, and is presently negotiating with the Australian National Library for the exchange at no cost of the records for imprints of the respective countries.

In the near term, LC will translate these records from the format of each country to a format resembling the Library of Congress MARC format as

closely as possible and make the records available through the MARC Distribution Serivice. The reason that we have to do this translation, a tailor-made program for each country's format, is that there does not exist today a true international MARC format. We hope that a year from now there will be such a format. Each national agency will then be able to have one translation program to convert from the international format to the national format and vice versa.

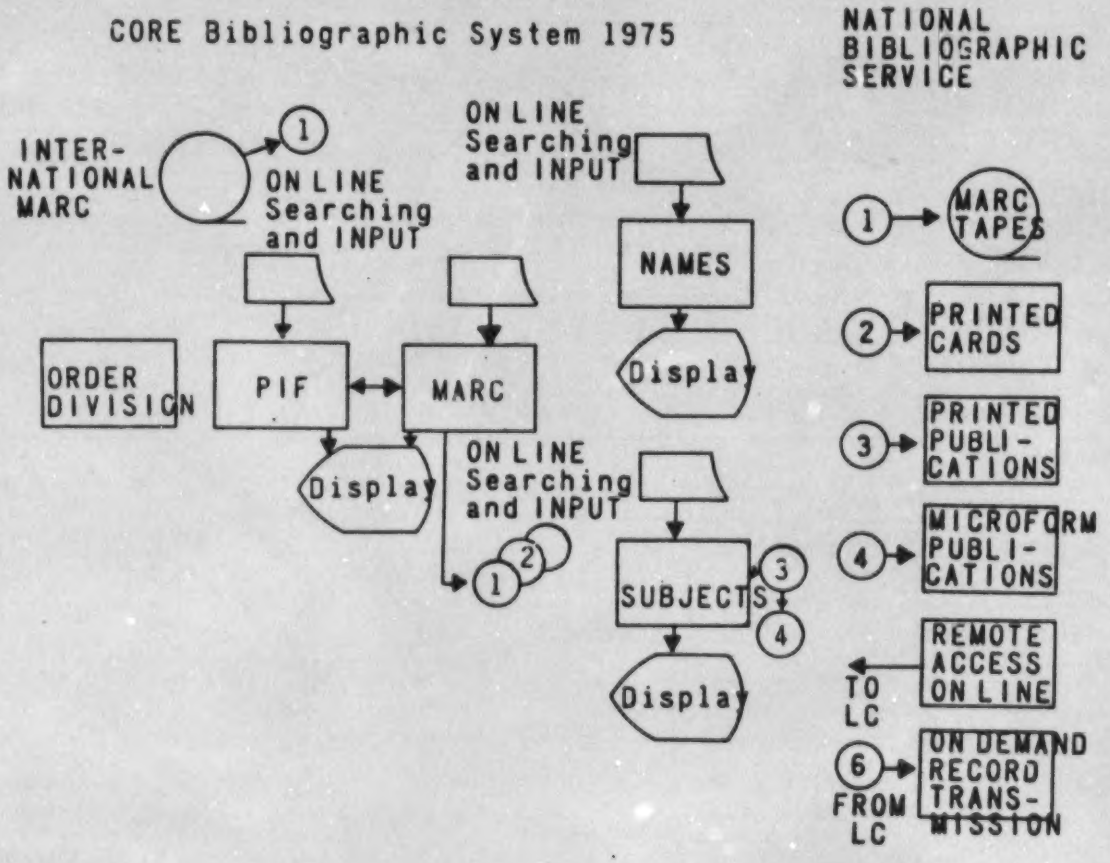
In 1976, it is planned to begin selecting records from the international tapes that the Library of Congress is going to process for inclusion in LC's Automated Process Information File. Those records selected will go through the LC processing stream and be distributed in the MARC Distribution Service as shown in Figure 33.

Figure 33.



The records not selected will also be maintained in machine-readable form and used by the Library of Congress as a residual file in the event that LC later needs to process one of the records. The international records will also be made available through the MARC service.

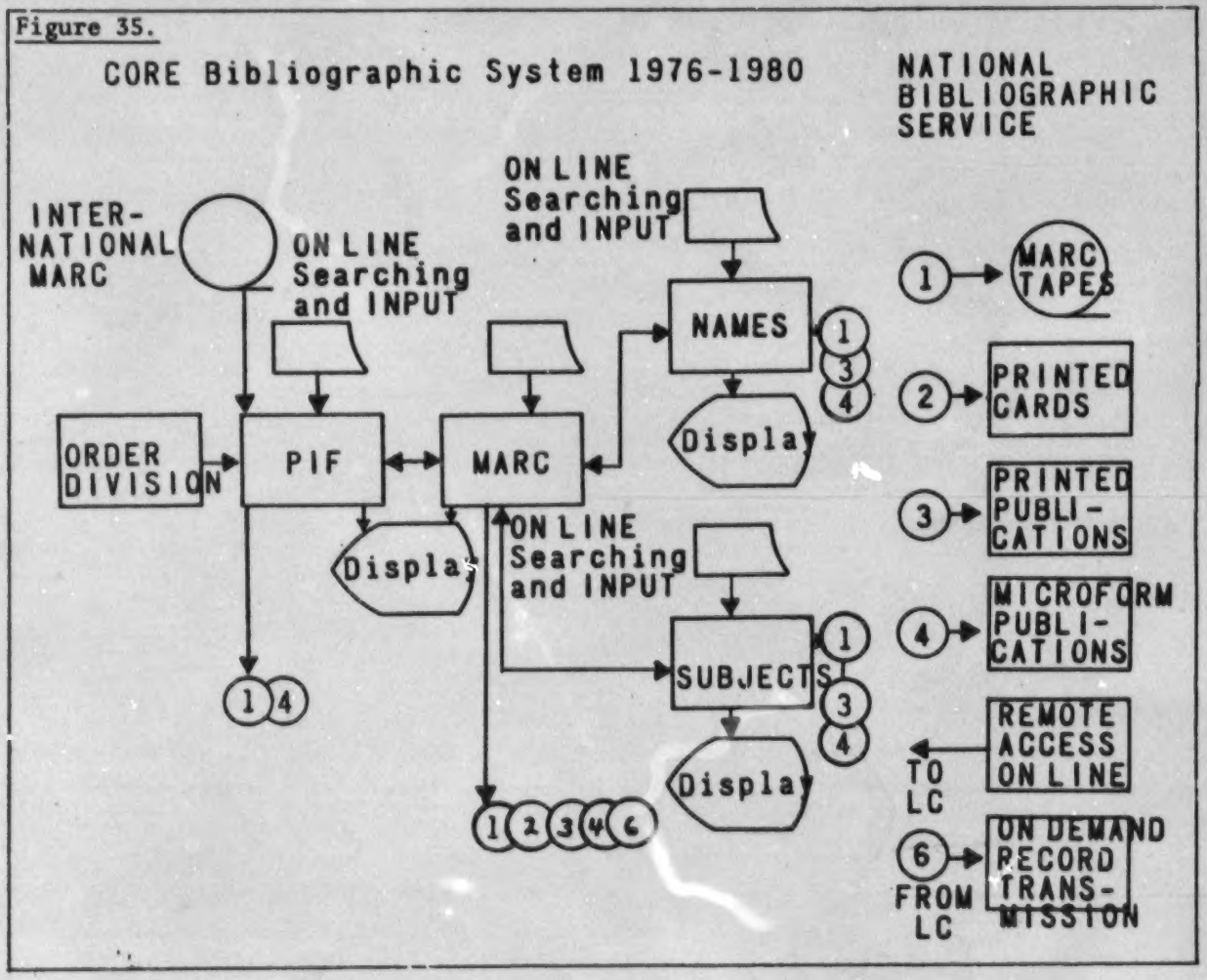
Figure 34.



Returning now to where Lucia Rather left off, I would like to summarize for you the Core Bibliographic System as it exists in 1975 (Figure 34). Behind all the projects that Lucia talked about, there was a plan, and this figure represents the plan. I said earlier that the function of the Core Bibliographic system is to put information under bibliographic control, and the approach is to build useful modules as we go along.

International MARC tapes are now coming in. The first tape that will be made available for distribution in this country will be the records from the National Library of Canada. The Order Division project is well on its way to completion. The Process Information File data will begin to be input in January 1976. MARC is now available online. We will also, in January, begin the online correction of MARC records. Subject headings are now available in printed form and in microform. We are working on the design of an online authority system.

Figure 35.

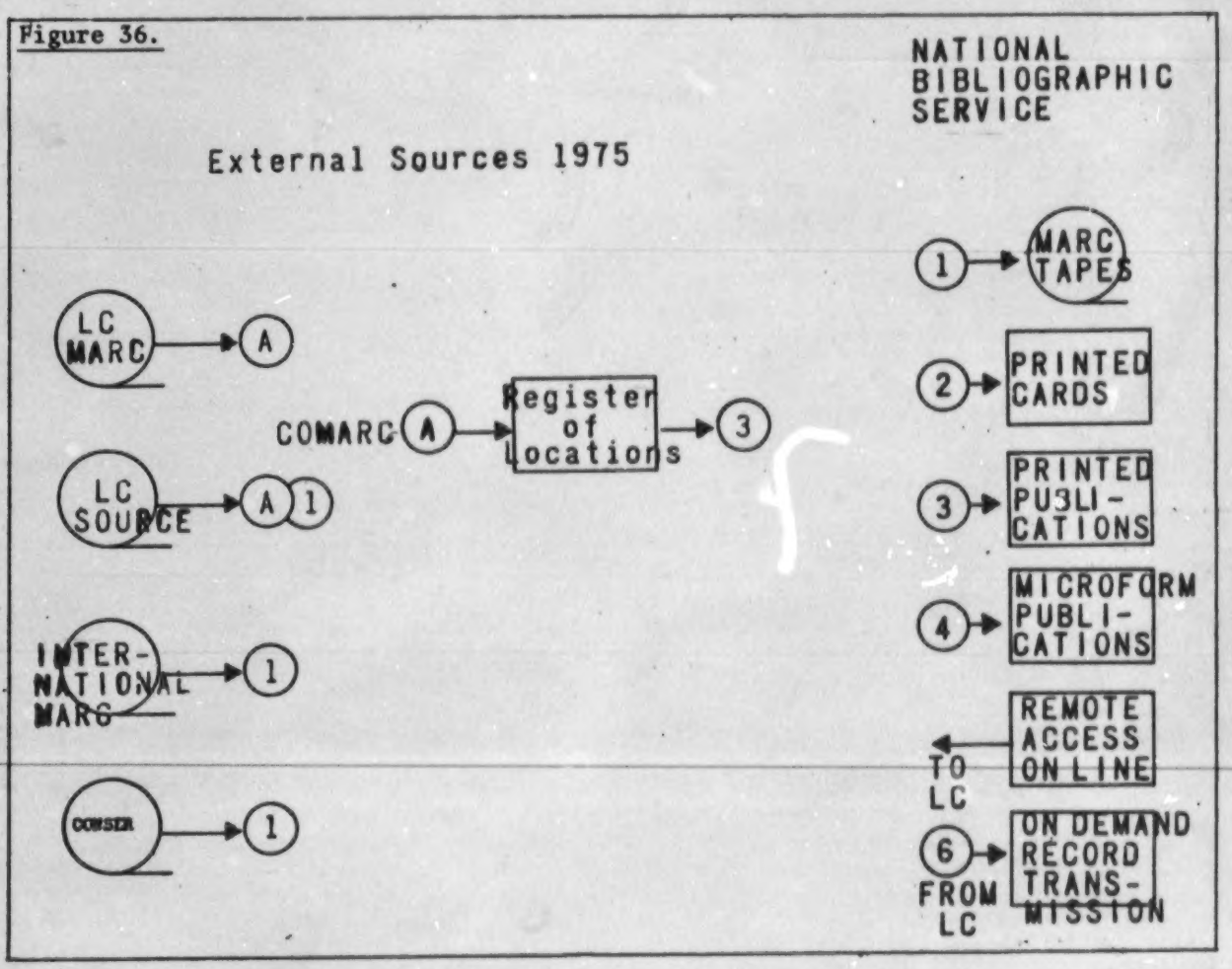


Our plans for the next four years are to begin to line these projects together as shown in Figure 35. Data from the Order Division will flow into the Process Information File. The Process Information File and the MARC file will be linked together. The international records will be fed into the Process Information File. MARC data will continue to be made available

as it is now, on tapes and as printed cards. MARC data will be the basis of LC printed publications and microform publications, and our plans include an on-demand record transmission service. Subject headings, in addition to being available in printed microform publications, will also be distributed in machine-readable form beginning in 1976 and name authority data will be made available in a variety of forms at a later date.

Although they do not appear on any of the handouts or transparencies, there are two other projects I wish to describe. The Research Libraries Group has proposed a joint project with LC which is not yet funded. The plan would be for a member of the Group to search the LC MARC data base for a desired record. If the record exists, it will be sent from the LC computer directly to the computer at the New York Public Library. Under this plan, the RLG would not maintain the entire MARC data base. This is the first time, to my knowledge, computer-to-computer transmission has been attempted in the library community, i.e., not for display, but a record transmitted for direct processing from a computer in one site to a computer in another site. Point 6 of Figure 35 refers to this concept of on-demand transmission. It is a very exciting prospect.

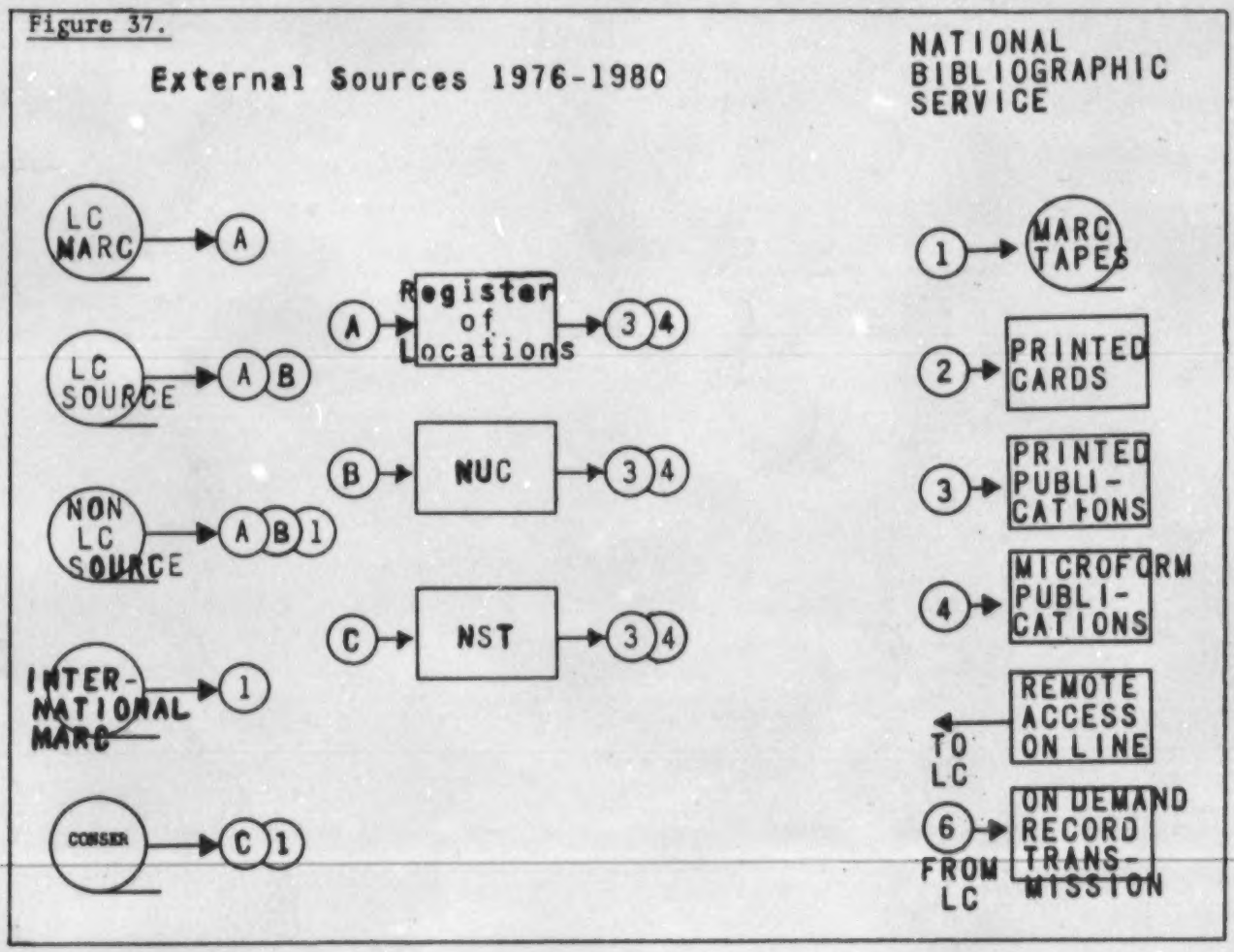
Figure 36.



At the same time, the Library of Congress is exploring the use of Tymshare. Several organizations have talked to us about online searching of the MARC data base at LC and we are now exploring using Tymshare to make this service available.

Figure 36 shown external sources. To recap, LC MARC records are being returned to LC for posting to the Register of Additional Locations and the Register is being made available in printed form. LC source records, under COMARC, are also being posted to the Register, and they will be made available through the MARC Service. International tapes are coming in, and they will also be made available through the MARC Service. And, of course, CONSER is imminent, and those records will be made available through the MARC Service.

Figure 37.



In Figure 37 we see the expansion of these external sources resulting in expanded national services from our machine-based files. The Library has recently submitted a proposal to a foundation to seek funds for a study involving these expanded services. The reporting format, the Register of Additional Locations project, the COMARC project, all taken together, are forming the nucleus of a national union data base in machine-readable form. The concepts are very complex. Following COMARC and the location reporting to the Register, we look to receiving nonLC source documents in machine-readable form. These, of course, will also be added to the Register and the national union data base and made available through the Distribution Service. Likewise, CONSER, in addition to becoming part of the MARC Service, will allow us to produce NST by automated means.

So we see the growth of the external sources. The Core Bibliographic System data and data from external sources, both national and international, taken all together, is what we have defined as the National Bibliographic Service, available in many forms--printed, microform, and online.

As part of all of this, the Library of Congress, under a contract award from the National Commission on Libraries and Information Science, is conducting a study with contractual support to define LC's role in the evolving National Bibliographic Network. The study has just begun, and the results should be very interesting to all of us.

* * * *

TRANSITION TO THE AUTOMATED SYSTEM

John C. Rather
Chief, Technical Processes Research Office
Library of Congress

MR. RATHER: Both Henriette and Lucia were too genteel and ladylike to tell you of the many problems that have to be faced in trying to bring this system into reality. We have bibliographic problems; we have technical problems on the computer side; we have procedural problems in trying to work a new system into a large ongoing operation; we have personnel problems in superabundance; we have political problems, and I will leave you to imagine what the ramifications of those may be. There are times when we are discussing some of the things that we have to do that Henriette and I try to console ourselves with a long-standing joke. We look at each other resignedly and say, "It can't be done." And it is at such moments that I find myself contemplating rather wistfully the pleasures of antiquarian books.

But the fact is that it can be and is being done. The work on the Core Bibliographic System, which will underlie the system that the Library of Congress will eventually use for its own bibliographic control, is proceeding and, though it may seem sometimes from the outside as if the pace is rather slow, I think the very deliberation of the pace is symptomatic of our desire to produce the best possible result.

The Library of Congress has traditionally been under many different pressures. We always have somebody on the outside telling us how we can best do our own work. And, though we try to be responsive to those needs, we still have to ensure the integrity of the internal processing system, which is indeed the core of our national service to libraries in this country. We are concerned with a number of problems as far as the implications of this system are concerned within the Library, primarily how we can best make the transition from the present manual system to the fully automated system that is being developed.

There are a great many other problems that have to be faced. Among them, of course, are the problems of ensuring that the new system preserves all of the essential values of the old system. We have under study or will have under study a number of problem areas. One that is being investigated at the present time relates to the inclusion of records containing nonroman languages in the MARC system. As you know, once you start talking about processing nonroman languages, particularly ideographic languages like the East Asian languages, you run into the character set problems that have been alluded to earlier. There are problems not only of inputting the characters and carrying them in the machine, but also of getting them out of the data base, to display them on CRT terminals, to

print them out, or to present them in photocomposed book catalogs. Each one of these tasks presents very definite technical problems but in trying to move ahead we hope that basic values that could be offered by the system will not be postponed indefinitely because we are hung up on some technical problem.

The Library now has a Working Group on Nonroman Languages in MARC. This is an interdepartmental group with representatives from the Reference Department, the Law Library, and the Processing Department. Even so, it is a relatively small group. The members collectively have an intimate and in some cases a native knowledge of the particular nonroman languages that we would like to add to the MARC data base. The kinds of questions we are exploring in this group are basically related to the advantages, if indeed there are advantages, of putting nonroman languages into the MARC system initially in a romanized form. Our feeling is that the expansion of the coverage of the MARC data base should proceed on a systematic basis, even though initially some languages -- Russian, Chinese, Japanese, the Indic languages -- may have to be entered in romanized form.

We recognize that there are definite disadvantages to such a decision. We are aware, however, that the disadvantages vary from language group to language group. One of the things that the working group is addressing is how satisfactory romanized records for particular languages would be in meeting the requirements of four different situations. The first question is: Does the system contain a known item? That is, if I know what I am looking for, will a romanized record answer the question? The second question is: If I have found a work that I know about, can I distinguish among editions from romanized records? The third question is: If I find the works of an author, can I determine what these works are from the romanized records? And the fourth question is: If I am looking under a subject heading and find several records, can I make a further distinction among them by reading the titles?

Those of you who are not familiar with the East Asiatic languages like Chinese and Japanese may find it surprising that even a native speaker of Japanese cannot always read a romanized form. The ambiguities of romanization in the ideographic languages create a real problem. Thus the answer to the fourth question is "Not very well." Indeed you have a progressive deterioration in the responsiveness of the system to each of those four requests.

So we recognize that, in dealing with this problem, we are going to have perhaps a kind of accommodation. Since there are many searches for known items, there is a definite advantage to having the romanized records in the data base even though you cannot answer all of the kinds of questions that you would like to have answered. We would try to resolve those more difficult problems by having some sort of backup file containing cards produced by our present methods.

A second problem of trying to make a transition to an operating situation where we rely completely on the machine data base relates to

shelflisting. Shelflisting is now done in relation to the totality of all of the records that we have ever entered into the system. If we come to rely on the MARC data base as the basis for our cataloging, we will have to modify the shelflisting procedures to make that possible. We do not want to be forever tied to having to look in the manual shelflist to determine whether we have arrived at a unique number. Shelflisting is an extremely costly process. It is one that we have to re-examine, simply because it is a form of close classification that currently costs us about $988,000 a year, a tidy sum that merits a little cost analysis.

We need to consider in more detail the nature of book catalogs that we will be producing for the library community, particularly as those book catalogs may come to be a kind of backup device for the system we are developing for use within the Library.

The register/index catalog seems to us to be an approach to catalog organization that would be beneficial. In that form of catalog, only the register contains the full information. There would be several indexes: one for names, one for titles, another for subjects. The information in any index entry would be complete enough to answer most of the questions that are asked in terms of getting at the materials. Those who seek cataloging data for some purpose would, of course, have to look in the register to get the full record. The connecting link would be a register number which would simply be sequential without special meaning. Thus the registers would never have to be restructured or cumulated.

A fourth area of concern is subject control. Mr. Welsh has mentioned that we do not anticipate ourselves being able to undertake a full-scale study of our subject heading system or of our classification system. But we are mindful of the fact that MARC records allow capabilities for subject retrieval that have not previously been available. For example, the geographical area codes in a fixed field contain encapsulated data that can be used to select records in an online retrieval system or in a batch processing system. It seems to us that this capability has definite implications for the nature of subject headings. There is discussion now about the possibility of double input of subject headings in the manual system to allow headings that say "Education-History-Indiana" to appear also in the form of "Indiana-Education-History." In a machine system you can access records through the geographical area code, so it is not necessary to have double input of the character string. The problems of storing indexes to subjects in a large system suggest also that there would be some advantages in examining a subject heading structure with a view to simplifying it for the purposes of computer file organization.

We have decided that, at that point when we begin to rely on the machine data base for current cataloging information, we will have to catalog in relation to that data base, and we cannot undertake to make the headings invariably compatible to those in the old card catalog. The implications of this have to be studied very carefully. We know in an operational sense why it is desirable. We want to be sure that the impact

is not a serious one. The whole area of authority files and their continuity and their character in the new system is an important one that must be studied carefully.

Finally, when we decide to close the card catalog, there is the question of the disposition of those catalogs. We will have to decide to what extent we can edit the catalog, in what forms can we afford to publish it, and what relationship that published form should have to the Mansell Catalog and the present book catalogs.

There are a number of factors that will ease the transition to the new system. If we decide to stop filing new cards in the card catalog about 1980 when all current records are going into MARC, we will already have a data base with one and one half million records; a million discrete names will be represented in it as will a couple of hundred thousand subject headings. And so a great deal of our present system will be preserved in the new form. We will not suddenly be cutting it off and turning our backs on it. We will have the advantages of the reference structure to provide the links between the machine data base and the old card catalog.

We have done some studies to examine the character of newly established names. We establish one new name heading for every two books we catalog. According to our studies, a surprisingly large proportion of those names are discrete as far as our present files are concerned, just on the basis of the information provided in the book. A much smaller proportion of the names require research to resolve conflicts than we previously thought.

Frequently when there is mention of closing the card catalogs, there is concern about serials. Plans for the CONSER project indicate that by 1980 there will be well over 200,000 serial records in machine readable form, and they will be the active serials that one would be concerned about. Some serial like the Brooklyn Chess Chronicle that expired in 1881 can safely repose in the old card catalog without any great loss to the system.

We are implementing the automated system in stages so that we have an opportunity to test and evaluate. We are very much concerned that not only the staff within the Library, but also the library community at large will be able to see the advantages of the course we are following well before it becomes fully operational.

In discussing what lies ahead with people in the Reference Department, you sometimes get the impression that they think some morning they will come in and the drawers of the Public Catalog will be nailed shut and they will be directed to a little TV screen and left to cope as best they can. Well, we intend a much smoother transition than that.

The implications of this new system are quite serious for the Library and for the library community. In the interests of trying to explore those implications, we are trying to maintain as high a profile as possible.

We are interacting with other groups. We have given many presentations such as this. I myself have talked about the future of catalog control at the Library of Congress so often, I feel like Harry Lauder making another farewell appearance.

We have an advisory committee on the future of the LC catalogs drawn from the membership of the technical services directors of large research libraries, and the six members of that group are giving us their reactions to our plans as we are able to formulate them.

We intend to publicize our plans as widely as possible to elicit reactions to them and to be as responsive as we can to those reactions, to the end that we will develop a system that benefits not only the Library of Congress but also the library community at large.

* * * *

Discussion

MR. WELSH: One general observation and then we will go on with the questions. I rise here in defense of something. We have been very critical collectively of the slow progress on CONSER. The progress has been slower than anticipated, but I would like to remind you that when this question of conversion of serials came up, ARL, the Library of Congress, and the Council on Library Resources met to decide upon the management of it. LC said that it could not do it, ARL said that it could not do it, and at that point Fred Cole volunteered to do it. So I think that before we criticize the Council on Library Resources, we ought to give Fred Cole and Larry Livingston a standing ovation for that initiative. Then you can criticize them.

MR. FORTH: I have found these sessions this morning very interesting. Are you all going to try and present this to other people?

MR. WELSH: We have no specific ideas about this. As you know, we have appeared before the RLG group; we have made a presentation to the Council for Computerized Library Networks; we made a proposal to the National Commission which expressed the very thought that you have just stated, that this is something that needs to be done. It is a twofold proposal: 1) we must have sufficient staff so that we can send some staff members out to tell the story, and 2) we must have the resources to bring people in to LC. We realize we should do it. I do not have any specific ideas at the moment. I know the MARC Institutes to some extent cover this.

MR. DE GENNARO: These proceedings are going to be published as part of the minutes of the meeting. At the January 1975 ARL Meeting when we had a program on the future of card catalogs, we published separately the presentations and this was quite successful. Perhaps this program could led itself to the same treatment; I can not guarantee it. We will have to see how it goes. But it has been very interesting and worthwhile, and we could and probably

should do something with it.

MR. HUMPHRY: I have a question relating to the Congressional Research Service and the possibility of accessing the data base that is generated for service. You may know that the New York State Library with a grant from the Council on Library Resources, is studying information services to state government. We have had conversations with the Library of Congress to determine whether it might be possible first to access the general non-confidential information in that data base, to enrich the service that is given to state government through the information that has been developed for CRS, and adapt that information to the state government requests. It could serve as a beginning search. It would not be used without some adaptation and caution.

The other part of the projection is that state governments are asking each other what information has been assembled on certain issues. In the long run it may be another network of state government requests that could be circulated among the state governments, with clearance with the Congressional Research Service.

I would hope that this might be a possibility for the future and another way in which the Library of Congress could serve, with its development of information, those agencies that need not do a lot of preliminary searching.

MR. WELSH: Normally I would say that is a policy question involving the Congressional Research Service, and I should not answer it, but I am not shy. With the caveats that you have attached to that, I think this is the very type of activity we are talking about.

The other half of that, from the processing point of view, is that I believe very firmly that LC should not be responsible for local and state publications. If each state assumed that responsibility, I think we could form a very sound working exchange on that basis.

MR. DE GENNARO: Could you clarify a little bit the RLG/LC connection? I think Henriette Avram said that it was in the state of not being funded yet. How certain is the funding? And how much funding is required? What is the timetable?

MR. WELSH: I think that I would prefer to ask James Skipper to respond to that.

MR. SKIPPER: This is a proposal for a pilot project which is now being formulated. In fact, the committee discussing it met this morning. The proposal is certainly not final; it is not approved by any authority, and it is not funded. The intent of Phase I of the proposal is to determine the economics and the benefits of one mode of access to a national data base. OCLC has already very successfully established a mode of utilizing this information. We are interested in exploring alternatives.

As Henriette Avram said, this will be on a limited basis, starting with NYPL and Columbia, with Yale and Harvard having the option to come in at some future date. It will be purposely a rather primitive communication link in order to hold down costs. We are fully committed to utilizing, adapting, and transferring existing software packages and technology without going into extensive development in this area.

Phase II will involve the more sophisticated technical complications of transmitting data between networks computer-to-computer rather than distributing catalog cards or putting it on a CRT.

MR. SHANK: Why is a COMARC format needed? Why won't the MARC format serve for the transfer of computographic data?

MS. AVRAM: We will be glad to send you the complete technical specifications. We had a very difficult time putting the COMARC rules on paper. We had enough data base formats given to us so that we recognized the variance from the MARC format. To send out such records as given would be very unfair to any recipient; none of their programs would work. It is too long a description to go through at this time, but the principal optional area that we are giving up in COMARC will be in the fixed-field area.

MR. KURTH: I wonder if Henriette Avram would remark on some of the principal criteria or ground rules for being a participant in COMARC.

MS AVRAM: The LC bibliographic source record would have to be converted as is. In other words, the data from top to bottom and left to right. Now there are a few fields in that record where this really is not that important, but principally it is the full bibliographic content.

As far as the tagging and the fixed fields and the remainder of the content designators, the fixed fields are almost all optional. There are some that are required purely for processing, like language. And, of course, we feel that that is an important field. The content designators for the variable fields, the bibliographic data, are just about what the MARC format is at the present time. There are some variations, but they are principally in the area of the fixed fields. If you would like us to, we would be glad to send you a set of specifications. The selection of participants has been done by receiving at LC their particular formats; and also it is a great help to us if we have test tapes that we can run through the system, because the whole idea of COMARC is to use the present software at LC. We could not have had a pilot project that quickly if we had had to write all new programs.

I would like to respond to the first question asked, about whether this presentation would be made again. This presentation will be made again in a great deal more detail by the Library of Congress Processing Department at the ISAD Institute in Washington, D. C. It will be a three-day meeting: two days of this in much greater detail, and then one day for tours at the LC.

MR. BOES: I would like to commend the Library of Congress and the Processing Department for the general direction of its plan. I would like to ask something which may be a little unfair, and that is: Has there been some discussion or movement in the direction of making the Processing Department and the Ohio College Library Center one, for obvious reasons?

MR. WELSH: I think in very, very serious terms that we are one. We view the OCLC as a very important link in the network. It is our position -- and this is a very strong position -- that we want to encourage OCLC to develop an authority system. We want to develop a system to make those authority files available to OCLC so that a consistent decentralized data base can be developed. And I think this should not be viewed as any form of competition. I think we want to cooperate. Just as Mr. Kilgour has been receiving the MARC tapes from the very beginning that form the basis for it, I think he will be receptive to receiving our authority file information, names and subjects, and together we can proceed to develop the national data base in a highly consistent and standardized form.

MR. BOES: I would obviously like to see a national system. That is why I like the general direction of your planning. It seems highly appropriate that OCLC is already moving in the direction of a national system, and this is a statement that the Library of Congress and OCLC should become one.

MR. KURTH: I just wanted to ask if Henriette Avram's concluding remark about this NCLIS study will properly address what Mr. Boes is mentioning.

MR. WELSH: It will. That reminds me that there is much developmental work taking place. I think you got some of the feel for it here this morning. The pace has quickened considerably. I hope that we can find a forum for continuous discussion. I think this is urgently needed, not only with you, but with the entire national and international library community. We are looking forward to the completion of the study that NCLIS has funded, plus some other studies that are under way which I think are going to give us some answers.

With reference to something that Warren Haas mentioned last night, there is a lack of enough planning to do the total job, and we are going to try to do something about that. I think we are much aware of the need for this, and of more communication from you.

For example, when I made the statement this morning about the evolving subject heading system, I was quite aware of how difficult it is going to be for one system to do all things for all people. I am aware that the Fine Arts people at New York Public Library will have some suggestions for improving the system, and, as I said this morning, we will welcome them. The Slavic specialists, and our Ukranian friends especially, have been very vocal on this, and we will try to work with them as well. We welcome this sort of interchange. We have got to be mindful that we are developing a system, though, that is going to meet the needs of everybody, and there are problems associated with that general approach.

MR. McDONALD: I wondered about the timing on CONSER. Is it on schedule? Do you believe that you will get it back into LC by 1977?

MR. WELSH: Yes. Henriette did mention that that was our expectation. It is not on the schedule that we originally conceived for a variety of reasons. The loading of both the MULS and LC file have been slower than planned. But this is a new venture. The management of it, bringing all of these people together to form a collaborative development, has taken a lot more time than anybody conceived. We hope that by November 1977 we will be able to take it back.

MR. GOVAN: I did not quite understand what Mrs. Rather had to say about the method by which the in-process information was to be captured and distributed to other libraries. Would you elaborate on that?

MR. WELSH: I can answer how it will be captured. In due course the Order Division and other records that are generated will be input to the in-process file. I think you have to appreciate that we are acquiring about 6 to 7 million pieces a year. And whereas we conceive that the Order Division's acquisitions, which number about a million pieces, can be input, unsearched, directly through the Order Division's automated system, the remainder will be input only after searching the file. So the Process Information File record, with its author, title and imprint information, begins at that point; it will be enriched at the final stage when subject headings and classification numbers are added. The intermediate record, as well as the final record, can be made available through the MARC Distribution Service, as well as -- and I think Henriette Avram made this point -- the cooperative work that is being done with the National Library of Canada, the Bibliothèque Nationale, and hopefully with the British Library, and with the National Library of Australia, getting those tapes, having them in our system, taking from them what we need to do your cataloging. The items that we do not select can also be distributed through the MARC Distribution Service.

MR. McDONALD: Can you tell us anything about the nature of the cooperation between LC and the A & I services in the creation of the national machine-readable bibliographic data base.

MR. WELSH: At the moment the problems involved in a cooperative effort are in the hands of the advisory group sponsored by NCLIS, NSF and CLR. Good progress has already been made, I believe. Two working groups have been established, one to create a MARC-like format for technical reports, and the other for journal articles. The advisory group is calling upon experts in those particular fields. For example, the chairman of one of the groups is Margaret Park; the chairman of the other group is Ann Curran. Those groups are being staffed by people who are involved in A & I-type activities, and the response has been fairly good so far.

MR. RATHER: Ann Curran's group is working on bibliographic name authority files. It is an interdisciplinary group that is exploring the possibility

of having common authority files in both the library sector and the A & I sector.

MR. WELSH: I would like, if I may, to make just one final remark. When John McDonald inquired if there was any possibility of the Processing Department presenting the program that you have had before you for the past two days, I said quickly, yes. We thank ARL for the invitation. And I want you to leave here with some feeling of what this meant to our staff. Each division of the Processing Department came up with an exhibit. At least 51 of you saw those exhibits. The staff was very much involved. They are aware of their clientele, and they were pleased. Thank you.

MR. DE GENNARO: I would like to have the last word here. I would like to thank William Welsh, the Library of Congress, and all the others for the tours that we had yesterday, for the excellent, and I would say, historic presentation today.

I think as Mr. Welsh said, the pace has quickened. It has not only quickened, but it has changed fundamentally. I think we are into a different period now in 1975. I think it is fundamentally different, and I think the pace is going to quicken even more. I am delighted that you all came to speak to us this morning. Thank you very much for that excellent presentation.

* * * *